FACTOR ANALYSIS OF GENETIC DIVERGENCE IN SESAME

BY

TES P. MATHEW

THESIS
SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENT FOR THE DEGREE
MASTER OF SCIENCE IN AGRICULTURAL STATISTICS
FACULTY OF AGRICULTURE
KERALA AGRICULTURAL UNIVERSITY

DEPARTMENT OF STATISTICS
COLLEGE OF VETERINARY AND ANIMAL SCIENCES
MANNUTHY - TRICHUR

1988

TO MY BELOVED
PARENTS

DECLARATION

I hereby declare that this thesis entitled "FACTOR ANALYSIS OF GENETIC DIVERGENCE IN SESAME" is a bonafide record of research work done by me during the course of research and that the thesis has not previously formed the basis for the award to me of any degree, diploma, associateship, fellowship, or other similar title, of any other University or Society.

TES P MATHEW

Mannuthy,
8-4-1988.

CERTIFICATE

Certified that this thesis, entitled "FACTOR ANALYSIS OF GENETIC DIVERGENCE IN SESAME" is a record of research work done independently by Kumari TES P MATHEW under my guidance and supervision and that it has not previously formed the basis for the award of any degree, fellowship, or associateship to her.

Saraswathi P.

Dr. (Mrs.) P. SARASWATHI
(Chairman, Advisory Board)
Associate Professor of
Agricultural Statistics
(NARP-SR)
College of Agriculture
Vellayani-Trivandrum

Vellayani,
8-4-1988.

APPROVED BY:

Chairman:

Dr. (Mrs.) P. Saraswathi Saraswathi P 24/6/88

Members:

Dr. K.C. George K. C. George 24/6/88

Mr P.V. Prabhakaran

Dr. Sverup John Sverup John

External Examiner

ACKNOWLEDGEMENTS

I am deeply indebted to Dr. (Mrs.) P. Saraswathi, Associate Professor of Agricultural Statistics, College of Agriculture, Vellayani, for suggesting the problem, expert guidance, continued support and encouragement rendered during the entire course of research work.

I owe my deep sense of gratitude to the members of my advisory committee Dr. K.C. George, Professor of Statistics, Mr. K.L. Sunny, Assistant Professor, College of Veterinary and Animal Sciences, Mannuthy, for providing computer facilities and correcting the manuscript and Dr. Sverup John, Assistant Professor of Plant Breeding, College of Agriculture, Vellayani for his consent to utilize his research data for my M.Sc. programme. I have great pleasure in recording my gratitude to my teachers.

The assistance rendered by Sri. K. Kalyan Raman, Research Officer in Statistics, Department of Statistics, University of Kerala, Karyavattom; Sri. C.E. Ajithkumar, Junior Programmer and Sri. C.T. Mukundhan, Technical Assistant, College of Agriculture, Vellayani, in computer programming and analysis is deeply acknowledged. I am thankful to Smt. K.P. Santha Bai, Junior Programmer;

Kumari Geetha, U., Department of Statistics, College of Veterinary and Animal Sciences, Mannuthy and to my friends who were helped me in so many ways at various stages of this study.

Grateful acknowledgement is made to Dr. K. Radhakrishnan, Dean-in-charge, College of Veterinary and Animal Sciences, Mannuthy and to the Kerala Agricultural University for the fellowship awarded to me during the course of study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

INTRODUCTION

INTRODUCTION

Multivariate techniques are generalizations of univariate techniques. Historically, statistical analysis on multivariate data was done on individual measurements by simpler univariate methods, which has got certain limitations in drawing overall inferences. Multivariate techniques helps to draw such overall inferences without loss of information we are seeking from the data.

Factor analysis is a branch of multivariate analysis developed by Spearman (1904) as a method of analysing the dependence structure of a set of variables. Today it is the most widely used method of reducing the dimensionality of a set of variables by taking advantage of their intercorrelations. This method helps to identify fundamental and meaningful dimensions of a multivariate domain. A matrix of correlations can be factorised in an infinite number of ways and a good account of these approaches are discussed by Harman (1967). Lawley (1940) applied the method of Maximum-Likelihood to estimate the loadings in a factor-model and now it remains the best method of extraction of factor loadings. The most commonly used technique is the principal factor analysis. But the estimates obtained by Maximum-Likelihood Factor

Analysis enjoys a powerful invariance property: "changes in the scales of the response variate only appear as scale changes of the loadings" (Morrison, 1978). In addition the maximum-likelihood method provides test of significance for the determination of the number of common factors. In a factor model response variate is represented as a linear function of a small number of unobservable common factor variates and a single latent specific variate. The common factor generates the correlations among the response variables while the specific factor contribute only to the variance of their particular responses.

A factor is a vector of correlation coefficients. The most interpretable factor is one based upon correlation coefficients which are maximally interpretable. The varimax criterion has become the most widely accepted method for the orthogonal rotation of the factors since its development by Kaiser (1958). In this method the variance of the squared loadings of each factor is maximised. The invariance property of normal varimax solution seem to be of greater significance and makes it to define mathematically the doctrine of simple structure. The rotation of the factors helps to make the best interpretation of the common factors.

In plant breeding programmes, a knowledge of the nature and magnitude of genetic diversity in morphological characters is important for careful selection of parents for crossing. The greater the diversity of genes that a breeder handles, better are the chances for the selection of superior genotypes. Correlation studies to evaluate the association of biometrical components on yield of sesame was conducted by many research workers. But no study has since been made to identify those hidden factors which have generated the dependent structure in the response variable. The present study is aimed at identifying those hidden factors by applying factor analysis via two methods - Principal Factor Analysis and Maximum-Likelihood Factor Analysis.

Sesame is one of the oldest annual oil seed crops. It has many favourable points, including a high percentage of oil which resembles, but in some respects superior to groundnut and sunflower oils. In respect of total world production, India stands next to China and India's share of world production is 24.6 percentage. Sesame is the most valued annual oil seed crop of Kerala. It is grown in very limited area of 1453 hectares with an annual seed production of 3648 tonnes. The chief factor limiting the productivity of sesame in the state is the lack of high yielding varieties suitable to the seasons in different regions. The only improved variety evolved in Kerala is Kayamkulam I. Sesame is grown in this state in uplands during rabi season and in wet lands during summer. Improved varieties suited to these varied conditions will step up the crop productivity in the state.

The biometric studies on genetic diversity on a large collection of different varieties of sesamum will provide basic information for its improvement in plant breeding programmes. In sesamum the germplasm diversity will affect the yielding ability mainly on the genetic characters that control them. The present study is conducted with the following objectives.

1. To create hypothesis on causative factors of divergence working in the plant population of sesame by means of 'exploratory factor analysis' through the method of "maximum likelihood solution".

2. To investigate the possibility of fixing fewer stable factors to delineate divergent plant populations.

3. To concentrate more on factors which are directly related to productivity, reproduction and vegetation.

4. To compare the results obtained through Principal Factor Analysis and Maximum Likelihood Factor Analysis.

5. To investigate the possibility of conducting a "confirmatory factor analysis" to test the hypothesis on factors and to find out the unique 'confactor position' to determine the unique position of factors.

REVIEW OF LITERATURE

REVIEW OF LITERATURE

Multivariate techniques are useful for analysing the intercorrelated multiple measurements. In plant breeding trials, as a large number of variables are involved, effective breeding calls for the knowledge of genetic variability among parents with regard to these characters which are sought to be improved. Genetic divergence among parents is important because a cross involving genetically diverse parents is likely to produce high heterotic effect and also more variability could be expected in the segregating generations. In such situation, factor analysis is an appropriate method which gives insight into the fewer causal influences (underlying factors) responsible for differentiation among genotypes or populations.

2.1 Theoretical studies

2.1.1 Analysis of dispersion

Attempts have been made to generalize the univariate analysis of variance to the case of multiple variates. The multivariate analysis of variance or MANOVA began with the derivation of the simultaneous sampling distribution of the variances and covariances in samples from a multivariate normal population (Wishart, 1928).

A few years later Hotelling (1931) found the distribution of a quantity T which is a natural extension of Student's t distribution to a sample from a multivariate normal population.

Wilks (1932 a), following the likelihood ratio method of Neyman and Pearson (1928, 1931) and Pearson and Neyman (1930) obtained suitable generalizations in the analysis of variance applicable to several variables and is the Λ statistic.

Bartlett (1934) applied it for testing significance of treatments with regard to two variables in a varietal trial and indicated its general use in multivariate tests of significance. Wilks (1935) and Hotelling (1936) found it useful for testing the independence of several groups of variates.

Bartlett (1947) demonstrated the useful approximation of Wilk's Λ statistic to a χ^2.

2.1.2 Factor analysis

Factor analysis is an exploratory tool to provide a posteriori insight into the underlying causes of correlations among variables in multivariate theory. An introspective analysis of the causative forces responsible for inter and intra-specific differentiation can be made by factor analysis. It provides fewer stable factors to delineate divergent populations.

Spearman (1904) developed first the theorems in factor analysis when he was attempting to understand the nature of intelligence as a single general factor among all tests of cognitive ability. Spearman differentiated three types of factors, namely, a general factor which was common to all of the variables, group factors which were common to some of the variables but not to all of them, and specific factors that were peculiar to single variables alone. In practice, the Spearman two-factor methods meet with the difficulty that group factors are frequently encountered. This two-factor theory was generalized in the next twenty years, principally by Garnett (1919) and Thurstone (1931), into principles of multiple factor analysis. The multiple factor method is supplementary to the Spearman's two-factor method in that there are no restrictions to the number of general factors or the number of group factors.

Roff's (1936) suggestion of filling each cell in the diagonal with the square of the multiple correlation of that variable with every other variable in the correlation matrix, had been advocated by Guttman (1956). It is the best known proposition and gives lower limit to the communality.

The computation on the correlation matrix has been divided into two basic methods of calculation which are in common use in extracting factors ie., in reducing a correlation matrix to a factor matrix. They are the principal axes method and centroid method. The latter was introduced by Thurstone as a substantial labour saving approximation to the principal axes method. Burt (1941) referred these methods respectively as 'the weighted summation' and 'the simple summation' methods. The centroid method of factor analysis was outlined by Holzinger and Harman (1941).

Thurstone (1947) traced the objective of the factor pattern as follows: "the object of a factor problem is to account for their inter correlations, in terms of a small number of derived variables, the smallest possible number that is consistent with acceptable residual errors".

Kendall (1950) made a useful distinction between dependence and interdependence analysis in multivariate analysis. Analysis of dependence is concerned with how a certain specified group depend on other and analysis of interdependence is concerned with how a group of variables are related among themselves. Factor analysis is latter type of multivariate analysis.

Burt (1952) has given a full amount of tests of significance in factor analysis developed upto that time.

The computation schemes of various factor analysis methods were provided by Frutcher (1954).

Rao (1955) made a distinction between the method of principal components and the common factor analysis. He introduced the concept of 'basis' of a vector space for the characterization of factor analysis. In the first characterization, the factor variable explains as much of variation as possible of original variables which lead to principal factor analysis. In the second characterization, he considered the factor variables which is predictable from the original variables with the maximum possible precision which lead to canonical factor analysis. According to the theory of canonical correlations the correlation (or its squares) between the two linear combinations of factor variables and test scores has to be maximized.

Factor analysis as a branch of multivariate analysis is very useful in determining the number and nature of causative influences responsible for the inter-correlation of variables in any population. Essentially, it aims at explaining a p x p correlation matrix (p variates) by means of a fewer number k ($k < p$) of meaningful factors (Maxwell, 1961; Lawely and Maxwell, 1963).

Kaiser and Caffrey (1965) explained the scale invariance aspects of canonical factor analysis and alpha factor analysis.

The better solution is to start by deciding the number of common factors and then allowing the communalities to adjust for it (Cattell, 1965 a). He has given a brief sketch of the whole process of factor analysis. He introduced the two types of factor models the 'closed model' for the method of 'component analysis' and the 'open model' for the 'factor analysis'.

In the closed model analysis unities are taken as the diagonal elements and take as many common factors as variables so that complete perfection as achieved within the small set of variables sampled.

The open model, using a reduced matrix with communalities in the diagonal, can produce fewer common factors than variables and this model will not enable to reinstate the test scores from the factor scores since full variance of any variable is partly contributed by the common factors and the rest by specific factors. He has given a good account of the approaches in deciding upon the number of common factors in a factor-model.

Cattell (1965 b) in another paper described the role of factor analysis in research. The use of factor analysis as (a) hypothesis creating and (b) hypothesis testing was given by him.

Hemmerle (1965) in his paper considered the problem of computing estimates of factor loadings, specific variances, and communalities for a factor analytic model. Iterative formulae were developed to solve the maximum likelihood equations and a simple efficient method of its implementation on a digital computer was described.

A general description of the concepts, theories and techniques of factor analysis has been given by Harman (1967).

Joreskog (1969) gave the relevant results for confirmatory factor analysis, where the matrix of factor loadings is uniquely identified by priori restrictions (usually by setting particular loadings to zero).

McDonald (1970) made a purely theoretical comparison among the three factor score construction methods namely principal factor analysis, canonical factor analysis and alpha factor analysis. According to him, in choosing a factor model, there are in fact, at least three separate choices to be made which are relatively independent. The first is the choice of basis in common factor space and it is the clearest defining characteristic of the three systems discussed. The second is the choice of an iterative algorithm for the determination of communalities/uniquenesses. The third is the decision rule for the number of common factors.

Joreskog (1971) has given estimation procedures for factor models involving several populations.

Joreskog and Gold berger (1972) have developed a generalized least-squares procedure. The estimates are scale free and asymptotically equivalent to the maximum likelihood estimates when the distribution is multivariate normal.

A non-metric approach to factor analysis has been considered by Kruskal and Shepard (1974). Although this technique has some attractive theoretical properties, it appears to be very sensitive to random variation in the data.

Swain (1975) considered a class of asymptotically efficient estimaters including both generalised least square and maximum likelihood as special cases and derived their large-sample properties.

Joreskog (1977) presents a general, all-encompasing series of methods for orthogonal factor analysis by the least squares and maximum likelihood methods. Many variables in the social sciences involve latent and structural variables and Joreskog (1977) developed estimation procedures for several such methods, working directly from the covariance matrix.

A few of the many methods developed for factor extraction are centroid method (Thurstone, 1947), principal factor method (Karl Pearson, 1901), maximum-likelihood method (Lawley, 1940) etc. Here we are considering principal factor and maximum-likelihood methods.

2.1.2.1 <u>Principal Factor Method</u> (PF method)

The literature on factor analysis contain a number of alternative methods and procedures for computation. Among these, principal component method (also called principal-factor or principal-axes solution) has several attractive features. Each factor or principal components as Hotelling calls it extracts maximum amount of variance and gives the smallest possible residuals. However, this method is preferred in the present study mainly owing to computational facilities.

Hotelling (1933 a) developed the principal axes method which provides an optimal solution at the suggestion of Kelley (1935).

Hotelling (1933 b) suggested the use of this method with either unities in the principal diagonal. The resulting factors are called "principal components" and are used to reproduce the score matrix rather than the correlation matrix. The number of principal components extracted is equal to the number of variables in the study.

Hotelling (1935) developed an iterative method of obtaining the loadings which can be carried to any degree of accuracy.

Principal component analysis is sometimes modified by the insertion of communalities in the diagonal of the correlation matrix and Rao (1955) called this method as principal factor analysis.

Harman (1960) exhibited an outline form of the numerical calculations of the method with an illustrative example. The first requirement in applying the principal factor method is to determine some suitable estimates of communality. According to him PF method can be considered as an excellent reduction of the correlation matrix which provides a basis for rotation to some other form of solution. The method also has the advantage of giving a mathematically unique (least squares) solution for a given correlation matrix.

Schilderinck (1978) has given a complete picture of the geometric and algebraic approaches of principal factor analysis.

2.1.2.2 Maximum-Likelihood Method (ML method)

The distinction between the solutions obtained by using the principal factor method and maximum likelihood method is that former corresponds to a priori choice of communalities and the latter, the number of common factors. The ML solution is based on fundamental statistical considerations. It considers explicitly the differences between the correlations among the observed variables and the hypothetical values in the universe from which they were sampled.

The efforts to provide a sound statistical basis for factor analysis were made first by Lawley (1940, 1942) who suggested the use of "maximum likelihood method", due to Fisher (1922, 1925), in order to estimate the universe values of the factor loadings from the given empirical data. Lawley's ML method is possible only when the variates are normally distributed. It requires a hypothesis regarding the number of common factors.

Lawley (1940) and Rao (1952) had shown that "ML solution" goes to and fro between communalities and number of factors until it hits on the combination which yields the smallest residual.

Kaiser (1960) recommended (after considering statistical significance, algebraically necessary conditions) the number of common factors as the number of eigen values greater than or equal to one in the correlation matrix. He found this number to be about one-sixth or one-third of the total number of variables. The expression of ML method in factor analysis becomes more meaningful and clear with this foundation.

A more condensed derivation of ML methods were appeared in a book by Lawley and Maxwell (1963).

Hemmerle (1965) found that Rao's procedure converges more rapidly than Lawley's procedure. Hemmerle (1965) in his paper considered the problem of computing estimates of factor loadings, specific variances and communalities for a factor analysis model. Iterative formulae were developed to solve the ML equations and a simple and efficient method of implementation of this method on a digital computer was developed by him.

The ML procedure remained impractical for all but for the smallest problems until the work of Joreskog (1967, 1969), as the process converge very slowly.

In Joreskog's (1967) ML method he proceeds systematically, fitting one, two, ..., factors and testing at each stage by a chi-square test to see whether further factors are required. It also carries a varimax rotation at each stage. He also presents an example to compare the ML factor estimates with those given by principal components.

Since Joreskog's work appeared, Clarke (1970) developed Newton-Raphson procedure for solving the log-likelihood function.

Lawley and Maxwell (1971) have given expressions for the standard errors of the latent roots and factor loadings for both the unrestricted and restricted models and later a correction in the standard error by Jennrich and Thayer (1973) was made.

Gill (1977) has shown that the ML estimators are consistent even when the underlying distribution for the variables is non-normal.

Kendall _et al_. (1983) reported that the ML solution remain scale-free if restrictions are imposed upon the parameters.

2.1.2.3 Factor rotation

Kaiser (1956) proposed the 'varimax' method as a modification of the quartimax method which nearly approximates simple structure. He found that a variable with communalities twice that of another will influence the rotations by four times as much.

As a last step in factor analysis Cattell (1965 a) explained the rotational technique like 1) simple structure and 2) confactor rotation. In simple structure each factor affects only a few variables. But in confactor rotation real factor does happen to operate on all or most of the variables in the sample.

Cattell and Khanna (1977) described different approaches to factor rotation in which he introduced one kind of rotation criterion i.e., confactor rotation, which arises when a second factorisation on the same variables with another group is involved.

2.2 Applied studies

Lawley (1943) applied the ML method to factor analysis of data collected for research in education. This is a satisfactory method of deciding the number factors required to account for the scores obtained when the number of individuals tested is reasonably large. In this case two general factors are needed to explain eight tests.

Wallace and Bader (1967) employed the multivariate approach of factor analysis on 27 measurements of the house mouse. The original principal factor solution was rotated to yield the final varimax solution. Five common factors were identified with respect to the 27 variables.

In order to determine the factors affecting the use of fertilizers among the farmers, Shetty (1969) used the principal component method of factor analysis. The study revealed that the first four factors are sufficient for the explanation of the observed inter-farm variations in the use of fertilizers.

Walton (1972) used factor analysis in identifying the morphological characters related to yield in spring wheats.

Abraham and Hoobakht (1974) applied the technique of factor analysis to extract basic factors underlying the observed soil variables. Scores based on four underlying factors could be used for comparison of inter soil variables.

Martin and Eaves (1977) adapted the analysis of covariance structures to the simultaneous maximum likelihood estimation of genetical and environmental factor loadings and specific variances. The goodness of fit is tested by chi-square and standard errors of parameter estimates can be obtained.

Denis and Adams (1978) performed a principal factor analysis on 22 morphological and yield-determining traits of 16 cultivars and strains of dry beans. There were at least two or three principal factors to be examined for biological meaning and from which to seek insight into the basic structural design of bean plants.

Tikka and Asawa (1978) used correlation in 28 genotypes of lentil for factor analysis through the principal component method as suggested by Harman. More than 90 percent of the variability was extracted by two factors. Within each factor, traits were ranked according to the relative magnitude of factor loadings.

Sundaram et al. (1980) used centroid method of factor analysis in cowpea to study its evolutionary pattern. The analysis divided the nine characters into three groups of factors which accounted for 98 percent of total variation.

Phenotypic correlations among seven traits in ninety diversified strains of triticale were utilized by Sawant et al. (1982) for factor analysis using the principal component method. The factor analysis grouped the seven characters into two main factors which together accounted for about 46 percent of total diversity.

Kendall et al. (1983) compared the ML factor estimates with those given by principal components by applying it to fifteen characteristics of 48 applicants for a post.

Kukadia et al. (1984) conducted a study to determine the importance of various traits for yield improvement in forage Sorghum. Genotypic correlations were subjected to factor analysis through the principal component method. Factors accounting for at least 10 percent variability were retained and arranged in order of variance.

Bartual et al. (1985) used factor analysis, principal component analysis and cluster analysis to identify sets of varieties better adaptable to the specific environmental conditions. Results obtained from ML factor analysis and principal component analysis were found to be similar.

A number of workers have discussed the importance of genetic diversity in plant breeding programme. A brief review of the past works are given as follows.

Murty and Anand (1966) brought out the role of genetic diversity in choosing parents for breeding programmes using a set of 10 varieties of linseed of diverse origin and F_1's.

Murty and Arunachalam (1967) have conducted a multivariate analysis of genetic divergence in the genus Sorghum (wild and cultivated types) using quantitative characters related to fitness under natural and human selection. They utilized factor analysis to compare the causal influences under natural and human selection for the diversity found in this genus. The factors were obtained by the centroid method. Factor analysis revealed the adequacy of the three factors for differentiation.

Multivariate analysis for measuring the degree of divergence between biological populations and for assessing the relative contribution of different characters to the divergence has been established by the contribution of several workers like Jeswani _et al_. (1970) and Somayajulu _et al_. (1970).

Ram and Panwar (1970) presented the result of multivariate analysis for a set of four characters related to productivity in 18 varieties of cultivated rice. The first two canonical roots accounts for 45 percent of the total variability.

Hussaini et al. (1977) studied the 640 genetic stocks of finger millet and the characters under study did not show high mutual correlation. Twelve groups have been identified by ploting the first two standardised principal components. Canonical variate analysis supported the above findings.

Gaur et al. (1978) studied the genetic divergence in 67 potato varieties/hybrids and found that the characters least influenced by the selection were mainly responsible for adding divergence to the population.

Singh et al. (1982) estimated genetic divergence among 48 exotic and 27 indigenous strains of chickpea.

Kamboj and Mani (1983) conducted a study to investigate the nature and quantum of diversity in a population of hexaploid triticales. Eight yield components for 100 genotypes were studied. The experiment was conducted in a simple lattice (repeated) design with four replications. Grain yield per plant and plant height contributed maximum towards genetic divergence.

On the basis of multivariate analysis, Valsalakumari et al. (1985) grouped 62 cultivars of banana into 8 clusters considering 22 characters simultaneously. The characters pulp/peel ratio on volume basis followed by weight of fruit contributed the maximum towards divergence.

Murugesan et al. (1979) assessed the heritability, co-efficient of variation, phenotypic and genetic advance as percentage of mean from 30 varieties of sesamum from the germplasm collections. The high heritability accomplished by high genetic advance indicated that most likely the heritability is due to additive gene effects and mass selection for such traits should be practiced.

Paramasivam and Prasad (1980) conducted a study of F_2 and F_3 populations of 3 crosses of sesame. They found that seed yield was positively and significantly associated with plant height, primary branches, secondary branches and capsule number. The above characters were also found to be associated among one another and showed the potentiality of these characters to be included in the selection programme.

A study was undertaken by Yadav et al. (1980) to find out the association of yield and its component characters in 22 genotypes of sesame.

Reddy (1981) noticed that the seed yield is a major component of oil yield in sesame. The seed yield can be improved by selecting the taller and less branched plants with high capsule number.

Reddy and Reddi (1984) made an investigation to assess the nature of gene action and to identify the better combiners for seed yield, oil yield and component characters in sesame.

Genetic parameters were assessed from nine varieties of sesame by Kandaswamy (1985). The results indicated the number of branches, number of capsule, number of seeds/capsule, and yield might be given due importance as indicated by additive gene effects in selection programmes as considerable improvement can be obtained in these characters.

Krishnadoss and Kadambavanasundaram (1986) studied the correlation of yield with six biometric characters in 125 varieties of sesame. Among these, three characters had high significant positive correlation with yield. They also had significant and positive intercorrelation among themselves. As such, improvement of these three important component characters will result in the improvement of yield in sesame.

MATERIALS AND METHODS

MATERIALS AND METHODS

3.1 Materials

The material consisted of biometric observations on 15 characters of 100 selected types of sesame varieties raised at the College of Agriculture, Vellayani during rabi (August to December) in 1981 and at the Kayamkulam rice research station during summer (January to April) in 1982. The varieties were grown in a simple lattice design, replicated twice. Each treatment consisted of 27 plants at 30 x 15 cm spacing with a plot size of 1.35 x 0.90m. A random sample of 10 plants per type per replication was selected and observations were recorded. The observations on the following characters were considered in the analysis.

1. Height of the plant (cm)
2. Number of branches
3. Height upto first capsule (cm)
4. Number of capsules on main stem
5. Number of capsules on branches
6. Number of capsules per plant
7. Number of fruiting nods/20 cm
8. Length of the capsule (cm)

9. Circumference of the capsule (cm)
10. Number of seeds/cap
11. Number of days to flowering
12. Number of days to maturity
13. 1000 seed weight (gm)
14. Seed oil content
15. Yield of seeds/plant (gm)

The varieties taken for the study are listed in Table 3.1.1

Table 3.1.1 List of sesame varieties taken for the study

Code Number	Name of the variety	Code Number	Name of the variety
1	Asthrango (local)	12	Gouri-Til
2	B14	13	Ie-284
3	B64	14	IS-20
4	BMI-3	15	IS-24
5	BM-5	16	IS-50
6	BM3-1	17	IS-47-GP-37-1
7	BM3-7	18	C-47
8	BS5-18-6(B)	19	Kayamkulam-1
9	Culture 7-1	20	Kayamkulam-2
10	ES-8	21	KRRI

Code Number	Name of the variety	Code Number	Name of the variety
11	G-P-111-2	22	KRR2
23	Muthukulam local No. 1	50	Si-44
24	Muthukulam No. 9	51	Si-47
25	Ie-34337	52	Si-52
26	IS-428-B	53	Si-57
27	Mutant K1	54	Si-71
28	North Kerala local No. 2	55	Si-72
29	" " " No. 3	56	Si-73
30	" " " No. 8	57	Si-99
31	" " " No. 17	58	Si-256
32	" " " No. 24	59	Si-259
33	No. 42	60	Si-261
34	NP.63	61	Si-267
35	KIS 300	62	Si-865
36	P10	63	Si-866
37	P16	64	Si-877
38	P23	65	Si-902
39	Ambalavayal	66	Si-914
40	P28 II	67	Si-2884
41	P38	68	Si-918
42	Si 1121	69	Si-925

Code Number	Name of the variety	Code Number	Name of the variety
43	Pt 58-35-1	70	Si-928
44	M3-1	71	Si-934
45	RT(1)	72	Si-950
46	RT(1)37	73	Si-951
47	Si-1	74	Si-953
48	Si-42	75	Si-973
49	Si-43	76	Si-976
77	Si-983	89	T12
78	Si-1001	90	T13
79	10-50-25	91	Si-2934
80	Si-1036	92	T177
81	Si-2834	93	Te-25
82	Si-1060	94	Sundar local
83	Si-1140	95	Timbl-9
84	Si-1141	96	TMV-1
85	Si-1154	97	TMV-3
86	Si-1159	98	Vayalathur
87	Si-1275	99	UT-43
88	S.Variety	100	Vinayak

3.2 Methodology

3.2.1 Structure of multivariate observations

Multivariate analysis is concerned with analysing multiple measurements that have been made on one or several samples of individuals. The multivariate analysis is concerned with the jointness of p measures on N subjects.

The mathematical model on which most of the multivariate procedures are based is on the assumption of multivariate normal distribution (m.n.d.). This assumption of m.n.d. for multiple measures can be justified by the same central limit theorem argument that leads to the assumption of normality for a univariate measurement. "The multivariate normal distribution often occurs because the multiple measurements are sums of small independent effects" (Anderson, 1958).

Measurements on p biometrical characters for N ($= n^2$) varieties replicated q times were denoted by $X_{ij\alpha}$ where ($i = 1,2,\ldots p$; $j = 1,2,\ldots q$; $\alpha = 1,2,\ldots N$). Suppose the random variables X_i of interest have a multivariate normal distribution with mean $\underset{\sim}{\mu}^{px1} = (\mu_1,\ldots,\mu_p)^1$

and covariance matrix $\Sigma^{p \times p} = (\sigma_{ij})$. If the measurements of interest are in widely different units, a more accurate picture of dependence pattern be obtained by standardising variable as $Z_i = \frac{X_i - \mu_i}{\sigma_i}$ $i = 1,2,\ldots,p$

Then analysis of the dependence structure of $Z_1,\ldots,Z_p$ which is given by the correlation matrix of $X_1,\ldots,X_p$, is done. Thus the observed correlation among variables constitute the original data.

3.2.2 Preliminary statistical analysis

The data were subjected to multivariate analysis of a simple lattice design. Lattice designs are an important class of row and column designs that are widely used in practice. Simple lattice design is an incomplete block design, and there are N (= n^2) treatments arranged in n blocks of size n replicated twice, the model to be fitted for the design is

$$\underline{X}_{jrs} = \underline{\mu} + \underline{\rho}_j + \underline{\beta}_{jr} + \underline{\tau}_{rs} + \underline{e}_{jrs} ;$$

$r,s = 1,2,\ldots\ldots,n$ (Federer, 1955)

Here all the above vectors consist of p elements corresponding to p characters of each variety.

where

μ is general mean effect

ρ_j is replicate effect in j^{th} replicate

β_{jr} is incomplete block effect in the j^{th} replicate

τ_{rs} is treatment effect

e_{jrs} is the random component

The analysis of variance for simple lattice design is given in Table 3.2.1.

Table 3.2.1 ANOVA for SLD

Sources of vatiation	d.f.	M.S.
Replication	$q-1$	
Treat (unadj.)	n^2-1	
Blocks within repln. (adj.)	$q(n-1)$	E_b
Intra block error	$(n-1)(qn-n-1)$	E_e
Treatment (adj.)	n^2-1	
Total	qn^2-1	

The treatments are adjusted by a weighting factor 'γ'

where

$$\mu = \frac{(E_b - E_e)}{n(q-1)E_b} \quad \text{if } E_b > E_e$$

$$= o \qquad \text{otherwise}$$

There is no need for adjusting the treatmental effects if $\mu = o$. The intra block error is used to test the significance of treatment effects (Cochran & Cox, 1957).

3.2.3 Analysis of dispersion

Multivariate analysis of variance was first developed by Wilks (1932 a). Analysis of dispersion is the process which involves the technique of analysing the variances and covariances of variables in multivariate case (Rao, 1952). The total dispersion is split up into various components as follows.

Table 3.2.2 MANOVA of p variables

Source	d.f	Dispersion matrix
Deviation from hypothesis	n^2-1	B
Error	$n^2(p-1)$	W
Total	pn^2-1	

The criterion arrived at by Wilks (1932a) through the generalised likelihood ratio principle is given by

$$\Lambda = \frac{|W|}{|W + B|}$$

where

W is the within dispersion matrix

B is the between dispersion matrix

The statistic used for testing the homogeneity of treatment means for all the characters taken together is given by $V = -m \log_e \Lambda$

where

V is distributed as χ^2 with (N-1) p degrees of freedom and $m = Nq - 1 + (p+N)/2$ (Bartlett, 1947).

3.2.4 Estimation of correlation matrix

The phenotypic, environment and genotypic correlations were estimated from the following analysis of variance-covariance of the data.

Table 3.2.3 Analysis of covariance of SLD

Source	d.f.	$MS(X_i)$	$MS(X_j)$	$MSP(X_iX_j)$
Replication	q-1			
Blocks (adj.)	q(n-1)			
Treatment (adj.)	n^2-1	MSV_i	MSV_j	MSV_{ij}
Intra-block error	(n-1)(qn-n-1)	MSE_i	MSE_j	MSE_{ij}
Total	qn^2-1	MSP_i	MSP_j	MSP_{ij}

Phenotypic correlation coefficient $r_{p_ip_j}$ =

$$r_{p_ip_j} = \frac{MSP_{ij}}{(MSP_i\ MSP_j)^{1/2}}$$

$$i \neq j$$

The environmental correlation coefficient

$$re_ie_j = \frac{MSE_{ij}}{(MSE_i\ MSE_j)^{1/2}} \quad i \neq j$$

&

The genotypic correlation coefficient

$$r_{g_ig_j} = \frac{(MSV_{ij} - MSE_{ij})/q}{\left[\frac{(MSV_i - MSE_i)}{q}\ \frac{(MSV_j - MSE_j)}{q}\right]^{1/2}} \quad i \neq j$$

3.2.5 Factor analysis

Factor analysis is the common term for a number of statistical techniques for the resolution of a set of variables in terms of a small number of hypothetical variables, called factors. It reduces the multiplicity of tests and measures to greater simplicity. The fundamental step in the analysis of a body of observed data is the formulation of a theoretical statistical model.

A linear model is used in order to explain observed phenomena in terms of simple theories. The following are the linear models employed in factor analysis.

3.2.5.1 Factor analysis models

Principal component analysis (Pearson, 1901 and Hotelling, 1933 a) and factor analysis (Spearman, 1904) are the two methods with different aims in analysing the structure of a covariance or correlation matrix.

Principal component analysis (PCA) is the method of reduction of a large body of data so that maximum of the variance is extracted. In this analysis, a set of p standardised variates $Z_1, \ldots\ldots, Z_p$ is transformed linearly and orthogonally into an equal number of new uncorrelated variables $F_1, F_2, \ldots\ldots F_p$. These are chosen such that F_1 has maximum variance, F_2 has the next maximum variance subject to being uncorrelated with F_1 and so on. The new variates are obtained by finding the latent roots and vectors of the correlation matrix. The linear model for component analysis is given by

$$\underline{Z} = \underset{\sim}{A}\,\underline{F} \qquad (1)$$

where

$\underline{Z}$ is the (px1) vector of Standardised Variables

$\underset{\sim}{A}$ is the pxp matrix of component loadings

$\underline{F}$ is the px1 vector of common factors

But the variance-oriented principal component analysis is not appropriate for investigating the correlation structure of the observed data, since all the components are needed to reproduce accurately the correlation coefficients among variables. However, for the application of the method no hypothesis need be made about the original variable.

In contrast to the method of principal component analysis, factor analysis is correlation-oriented and explains observed correlations among variables in terms of smaller number of hypothetical factors.

The basic factor analysis model can be written in matrix notation as

$$\underline{Z} = \underset{\sim}{A}\ \underline{F} + \underline{e} \qquad (2)$$

where

$\underline{Z}$ is the px1 vector of standardised variables

$\underset{\sim}{A}$ is the pxk matrix of factor coefficients

$\underline{F}$ is the kx1 vector of $k < p$ common factors

$\underline{e}$ is the px1 vector of specific (unique) factors

This equation states that the observed variables are weighted combinations of the common factors and the unique factors. The common factors account for the correlations among the variables and the unique factor account for the remaining variance including error of that variable. The total unit variance of a standardised variable Z_i is made up of the communality attributable to the common factor and the uniqueness, which is the contribution of the unique factor (Harman, 1967).

In factor analysis it is usual to discard the sample mean vector and to make use of the covariance matrix or correlation matrix alone. The dispersion matrix of the variates in $\underline{Z}$ is defined as $E\ (ZZ')$ and is symmetric and positive definite of order p. The assumptions are

$$E\ (Fe') = 0 \qquad (3)$$

$$E\ (FF') = I_K \qquad (4)$$

$$\&\ E\ (ee') = \psi \qquad (5)$$

where ψ is a diagonal matrix with diagonal elements as ψ_i

Since $E\ (ZZ') = E\ \left[(AF + e)\ (AF + e)'\right]$

we have $R = AA' + \psi \quad (6)$

In practice A and ψ are unknown parameters which are to be estimated from experimental data.

Principal factor analysis method, centroid method, maximum likelihood method, minimum residual method etc. are some of the methods for estimating the parameters A and ψ . Among these methods some require estimates of communalities while others require estimates of the number of common factors.

3.2.5.2 Exploratory versus confirmatory factor analysis

A particular application of factor analysis is exploratory or confirmatory according as the number of parameters prespecified in the model equation of factor analysis (Joreskog, 1969). In this study exploratory factor analysis is done by the principal factor analysis and maximum likelihood methods.

3.2.5.3 Estimation of communality

Communality is the amount of variance of the characters accounted for by the common factors (Frutcher, 1954).

There are various methods of estimating communality. But the squared multiple correlation (SMC) of each variable with all other variables of the set seems to be the 'Best

Possible' systematic estimate of communality (Guttman, 1956).

The SMC of variable Z_i is given by

$$SMC_i = R^2_{i.1\ 2\ \ldots..(i-1),\ (i+1)\ldots..p} = 1 - \frac{1}{r^{ii}} \quad (7)$$

where r^{ii} is the diagonal element of R^{-1} corresponding to the variable Z_i. The SMC has another important property that it is the lower bound of the communality (Harman, 1967).

The maximum correlations in corresponding row or column may also be taken as initial estimates of communality (Cattell, 1965 a).

3.2.5.4 <u>Principal factor analysis</u> (PFA)

The application of the principal components to the reduced correlation matrix with estimates of communalities in the diagonal instead of ~~ones~~ unities leads to the principal factor analysis. This method yields a mathematically unique solution of component correlations.

From the classical factor, analysis model (2) the relevant portion of the determination of the common factor coefficients may be

$$\underline{Z} = \underset{\sim}{A}\ \underline{F} \tag{8}$$

or

$$\left.\begin{aligned} Z_1 &= a_{11}f_1 + \ldots\ldots + a_{1k}f_k \\ &\cdots\cdots\cdots\cdots \\ &\cdots\cdots\cdots\cdots \\ Z_p &= a_{p1}\,f_1 + \ldots\ldots\ldots + a_{pk}\,f_k \end{aligned}\right| \tag{9}$$

The sum of squares of factor coefficients gives the communality of a particular variable while $a_{im}^{\ 2}$ indicates the contribution of the factor f_m to the communality of Z_i. The principal factor method involves the selection of the first factor coefficients a_{i1} so as to make the sum of the contribution of that factor to the total communality a maximum.

ie, $V_1 = a_{11}^{\ 2} + \ldots\ldots + a_{p1}^{\ 2}$ (10)

is maximum. The coefficients a_{i1} must be chosen such that V_1 is maximum under conditions.

$$r_{ij} = \sum_{m=1}^{K} a_{im}\,a_{jm} \tag{11}$$

$$i, j = 1, 2, \ldots\ldots p$$

where

$r_{ij} = r_{ji}$ and r_{ii} is the communality $h_i^{\ 2}$ of the i-th variable.

This condition implies that the observed correlations are to be replaced by the reproduced correlations, implying the assumption of zero residuals. V_1 is maximised by applying the method of lagrangian multipliers under the conditions (11). The maximisation of V_1 leads to the system of p equations in p unknown a_{i1}

$$\text{ie, } (R_1^* - \lambda_1 I)\, q_1 = 0 \qquad (12)$$

where R_1^* is the reduced correlation matrix

$$\text{ie, } R_1^* = R_1 - \Psi$$

q_1 is the latent vector corresponding to the latent root λ_1

$$\lambda_1 = \sum_{i=1}^{p} a_{i1}^2 \text{ and } \lambda_2 = \sum_{i=1}^{p} a_{i2}^2 \text{ and so on.}$$

The linear homogeneous equation system (12) has only a non-trivial solution if its determinant is equal to zero.

$$\text{ie, } |R_1^* - \lambda_1 I| = 0 \qquad (13)$$

The criterion regarding the number of common factors to retain in the factor model is equal to the number of principal components whose eigen values are greater than one. The investigator will usually be satisfied with an even smaller number of factors.

The characteristic equation (13) gives latent roots $\lambda_1, \lambda_2, \ldots\ldots \lambda_k \geq 0$ and the associated orthogonal characteristic vectors $\underline{q_1}, \underline{q_2}, \ldots\ldots, \underline{q_k}$

Jacobi method is used to find out the eigen values and vectors of the matrix A. The idea of the Jacobi's method is to pick up the largest off-diagonal element of the matrix and to 'annihilate' it to zero by applying a proper orthogonal transformation. Then the largest remaining off-diagonal element is found out and that is annihilated. The procedure is repeated until the off-diagonal elements were sufficiently close to zero or negligible. The diagonal elements of the matrix is a close approximation to the eigen values. If the successive transformation matrices were multiplied together, they would produce an accurate approximation to the matrix of eigen vectors (Mulaik, 1972).

Substituting the largest characteristic root λ_1 in (12) we get corresponding characteristic vector.

$$\underline{q_1}' = (q_{11}, q_{21}, \ldots\ldots q_{p1}) \tag{14}$$

The normalized characteristic vector $\underline{\hat{q}_1}$ which fulfil the conditions (10) and (11) is

$$\underline{\hat{q}_1} = \frac{\underline{q_1}}{(\underline{q_1}' \cdot \underline{q_1})^{1/2}} \tag{15}$$

then the first column vector of factor loading matrix is determined as $a_1 = \underline{\hat{q}_1}\sqrt{\lambda_1}$ (16)

The second column vector of A is $a_2 = \underline{\hat{q}_2}\sqrt{\lambda_2}$ and so on. This shows that $a_1, a_2, \ldots\ldots$ are scaled normalized characteristic vectors.

The sum of the squares of factor loadings of the variable gives the corresponding communality ie, the squared factor coefficients can be considered as the percentage variance components of the common factor (Harman, 1967). The iteration process is continued with the new estimates of communalities until a specified degree of convergence is occurred. The controlling equation to ensure that no vital information is lost is

$$R_1^* = AA' \tag{17}$$

There are many equivalent matrices which all satisfy $R_1^* = AA'$. It implies also the making of a reasonable choice among the many possibilities to perform a final matrix A, which contains a suitable interpretation of the relation under research. This results in the rotation of the factors of the initial matrix A.

3.2.5.5 Factor rotation

After extraction, the matrix of factor loadings are submitted to varimax orthogonal rotation, the effect of which is to accentuate the larger loadings in each factor and suppress the minor loading coefficients, and in this way improve the opportunity of achieving a meaningful biological interpretation of each factor (Denis and Adams, 1978).

Kaiser's (1958) varimax rotation is one in which factors are rotated in such a way that the new loadings tend to be either relatively large or relatively small in absolute magnitude compared with the original ones. The simplicity of a factor is defined as the variance of its squared loadings

$$V_k = \frac{\left\{p \sum_{i=1}^{p} (a_{im}^2/ h_i^2)^2 - (\sum_{i=1}^{p} a_{im}^2/ h_i^2)^2\right\}}{p^2} \qquad (18)$$

where a_{im} is the new factor loading for variable i on factor m, where $i = 1,2\ldots\ldots p$ & $m = 1,2\ldots\ldots K$

For entire factor matrix the normalized varimax criterion is

$$V = \sum_{m=1}^{k} \frac{\left[p \sum_{i=1}^{p} (a_{im}^2/h_i^2)^2 - (\sum_{i=1}^{p} a_{im}^2/h_i^2)^2\right]}{p^2} \qquad (19)$$

where h_i^2 is communality of i^{th} variable.

The fundamental rationale for attempting to establish the normal varimax criterion is that the normal varimax solution is invariant under changes in the composition of the variables.

3.2.5.6 Maximum likelihood factor analysis

Alternate methods that circumvent many of the problems of principal factor analysis have been suggested. One such method is maximum-likelihood factor analysis proposed by Lawley (1940) and later which provides maximum likelihood estimates for the factor loadings. Maximum likelihood solution requires an estimate of the number of common factors. A ML solution has the same general appearance as a PF solution, but it does not have the latter's property of accounting for a maximum amount of variance for a specified number of factors. Also, while a PF solution is unique for a given body of data, a ML solution differs from another by a rotation (Harman, 1967). When estimating a population parameter, if a sufficient statistic exists to estimate the parameter, the maximum likelihood estimator is usually based on it. Moreover, the ML estimator is a consistent estimator as well as frequently a minimum variance estimator (Mulaik, 1972). A well known property of ML method of factor analysis is that it is independent of the units of measurement in the characters.

The model to be used in this method is (2). Also X follows multivariate normal distribution with mean vector μ and covariance matrix Σ.

The sample covariance matrix of X is denoted by S where $S = \frac{1}{n} \sum_{\alpha} (X_\alpha - \bar{X})(X_\alpha - \bar{X})'$

$$\bar{X} = \frac{1}{N} \sum_{\alpha=1}^{N} X_\alpha$$

X_α is the column vector of random sample of N ($>$ p) observations of X. $\alpha = 1, 2, \ldots\ldots N$, $n = N-1$

The distribution of S is Wishart with n d.f. ie, $ns \sim W(\Sigma, n)$

Here $E(S) = \Sigma$

The logarithm of the likelihood function for the sample, omitting a function of the observations, is given by

$$\log_e L = -\tfrac{1}{2} n \left[\log_e |\Sigma| + tr \; (S \Sigma^{-1}) \right] \qquad (20)$$

This is regarded as a function of A and ψ. Considering these as mathematical variables we seek values of A and ψ denoted eventually by $\hat{A}$ and $\hat{\psi}$ that maximize the value of $\log_e L$. It is more convenient to minimize the function.

$$F_K(A, \psi) = \log |\Sigma| + tr \; (S \Sigma^{-1}) - \log |S| - P \qquad (21)$$

For the purpose of minimising the function F the partial derivatives with respect to the elements of A and the diagonal elements of ψ which is given by

$$\frac{\partial F}{\partial A} = 2\, \Sigma^{-1} \; (\Sigma - S)\, \Sigma^{-1} A \tag{22}$$

$$\frac{\partial F}{\partial \psi} = \operatorname{diag}\left[\Sigma^{-1} \; (\Sigma - S)\, \Sigma^{-1}\right] \tag{23}$$

are required.

Equating $\frac{\partial F}{\partial A}$ and $\frac{\partial F}{\partial \psi}$ to zero and solving the resulting equations to get the estimates of A and ψ (Lawley & Maxwell, 1971). The estimation equations are independent of the scale of measurement of the X's and consequently the estimation equations for the a's can be expressed in terms of the correlations rather than the covariances. Lawley (1940)

$$\text{ie, } R = AA' + \psi \tag{24}$$

$$\text{and } \psi = I - \operatorname{diag} AA' \tag{25}$$

$$A' R^{-1} A \text{ is diagonal} \tag{26}$$

premultiplying both sides of (24) by $A' \psi^{-1}$ yields

$$(A' \psi^{-1} A + I) A' = A' \psi^{-1} R \tag{27}$$

This equation can be simplified to

$$JA' = A' \psi^{-1} R - A' \tag{28}$$

where

$$J = A' \psi^{-1} A \tag{29}$$

which is amenable to an iterative method of solution (Lawley, 1942).

Starting with an arbitrary factor matrix $A = (\underline{a}_1, \underline{a}_2, \ldots\ldots, \underline{a}_m)$ (usually loadings obtained from principal component analysis) and corresponding

$$\psi_1 = I - \text{diag } A\,A' \tag{30}$$

the factor loadings $B = (\underline{b}_1, \underline{b}_2, \ldots\ldots, \underline{b}_m)$ are derived from the iterative process, where

$$\underline{b}_1 = (R\ \psi^{-1}a_1) \,/\, [a_1^1\ \psi^{-1}\ (R\ \psi^{-1}a_1 - a_1)]^{1/2}$$

$$\underline{b}_2 = (R\ \psi^{-1}a_2 - a_2 - b_1 b_1^1\ \psi^{-1}a_2) \,/$$

$$[a_2^1\ \psi^{-1}\ (R\ \psi^{-1}a_2 - a_2 - b_1 b_1^1\ \psi^{-1}a_2)]^{1/2}$$

$$\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot$$

$$\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot$$

$$\underline{b}_m = (R\ \psi^{-1}a_m - a_m - \ldots\ldots - b_{m-1}\ b^1_{m-1}\ \psi^{-1}a_m) \,/$$

$$[a_m^1\ \psi^{-1}(R\ \psi^{-1}a_m - a_m - \ldots\ldots - b_{m-1}\ b^1_{m-1}\ \psi^{-1}a_m)]^{1/2}$$

$$\psi 2 = I - \text{diag } BB^1$$

The iterative process is repeated again and again until the convergence is obtained to the desired degree of accuracy. In standardised variates, the convergence criterion has usually be taken as 0.005. The final matrix A contains the ML estimates of factor loadings for the assumed number of common factors. In this iterative method it is tacitly assumed that none of the uniquenesses vanish. In some cases the maximisation of the likelihood function leads to one or more of the variables with uniqueness essentially zero. In the literature of factor analysis this type of improper solutions have usually been known as Heywood cases. Joreskog has made a provision for the Heywood case.

It is assumed that a maximum likelihood factor analysis with a certain value of k has been performed resulting in an improper solution with m ($\leq k$) of the unique variances zero. Assuming that this has occurred for the first m variables, the dispersion matrix may be partitioned as

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \tag{31}$$

where

Matrices S_{11}, S_{12}, S_{21} and S_{22} are of orders m x m, m x (k-m), (p-m) x m and (p-m) x (k-m) respectively.

Then the estimates $\hat{A}_{11}$, $\hat{A}_{12}$ and $\hat{A}_{21}$ are defined as

$$\hat{A}_{11} = S_{11} \Gamma \; \Delta^{-1/2} \tag{32}$$

$$\hat{A}_{21} = S_{21} \Gamma \; \Delta^{-1/2} \tag{33}$$

$$\text{and } \hat{A}_{12} = 0) \tag{34}$$

where Γ is an orthogonal matrix of order mxm that reduces S_{11} to diagonal form and Δ is a diagonal matrix containing latent roots of S_{11}. The matrices $\hat{A}_{22}$ & $\hat{\Psi}_2$ are obtained by applying the maximum likelihood method to the conditional dispersion matrix

$$S_{22.1} = S_{22} - S_{21} \; S_{11}^{-1} \, S_{12} \tag{35}$$

In the analysis of $S_{22.1}$ the number of variables is decreased by m and also the number of factors is decreased by m. Then

$$\hat{A} = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix} \text{ and } \hat{\Psi} = \begin{bmatrix} 0 & 0 \\ 0 & \hat{\Psi}_2 \end{bmatrix}$$

are the maximum likelihood estimates of A and Ψ

3.2.5.6.1 <u>Test of significance for the number of factors</u>

One of the main advantages of using the maximum likelihood method of estimation is that it enables us to test the hypothesis H_k that, for specified k, there are k common factors. After obtaining a proper solution the hypothesis is tested by

$$U_k = \left[N-1 - (2p+5)/6 - 2k/3\right] f_k(\hat{\psi})$$

where

$$f_k(\hat{\psi}) = \sum_{i<j} \left[(S_{ij} - \hat{\sigma}_{ij})^2 / \hat{\psi}_i \; \hat{\psi}_j\right] \text{ and}$$

$S_{ij} - \hat{\sigma}_{ij}$ represents the residual covariance of x_i and x_j after eliminating k common factors. The criterion U_k is actually a measure of how much the residual covariances differ from zero. Under H_k, for moderately large n, U_k is very nearly distributed as χ^2 with d_k d.f.

where $d_k = \frac{1}{2}\left[(p-k)^2 - (p+k)\right]$

This exactly imposes an upper limit on m for given p. ie, The number of common factors cannot exceed the largest integer satisfying

$m < \frac{1}{2}(2p+1 - \sqrt{8p+1})$ for a fixed number of p variables.

The non significance of χ^2 means that there would be no point in fitting further factors to the data.

The computations were carried out on the VERSA IWS system in the statistics department of the KAU. The computer programmes used for the analysis are given in Appendix-VII.

RESULTS

RESULTS

The results obtained by the application of appropriate statistical techniques on the data generated from experiments conducted in the uplands (data A) and rice fallows (data B) are given below.

4.1 Results of data A (uplands)

4.1.1 Preliminary statistical analysis

The analysis of variance for simple lattice design was made for each character under study. The F-values for testing equality of each character are given in Table 4.1.1. All characters except the number of fruiting nodes per 20 cm were found to distinguish the genotypes. The mean values of the various characters are presented in Appendix I.

4.1.2 Analysis of dispersion

Multivariate analysis of variance was done and the total dispersion matrix was split up into 'between' and 'within' dispersion matrices. The between and within dispersion matrices are given in Appendices II and III respectively. The value obtained for Wilk's lambda statistic was

$$\Lambda = 7.2 \times 10^{-8}$$

So that V = 2327.33 which is distributed as a

Table 4.1.1 F-values obtained from the analysis of variance for 15 characters - data A

Sl. No.	Character	F-values
1.	Height of the plant	3.9557**
2.	No. of branches	5.6030**
3.	Height upto first capsule	3.9170**
4.	No. of capsules on main stem	3.2020**
5.	No. of capsules on branches	2.1530**
6.	No. of capsules/plant	2.3661**
7.	No. of fruiting nods per 20 cm	1.1764
8.	Length of capsule	1.7876**
9.	Circumference of capsule	2.8971**
10.	No. of seeds/capsule	2.6235**
11.	No. of days for flowering	6.0745**
12.	No. of days to maturity	2.1138**
13.	1000 seed weight	18.8237**
14.	Oil content	570.7724**
15.	Yield of seeds/plant	3.8164**

** Significant at 1% level

* Significant at 5% level

chi-square with 1485 degrees of freedom and is significant at one percent level of significance.

4.1.3 Estimated correlation matrices

The analysis of covariance was done for all the combinations of 15 characters under study. The phenotypic, environment and genotypic correlation coefficients for each pair of characters were calculated from the corresponding variance-covariance components and are given respectively in Tables 4.1.2, 4.1.3 and 4.1.4. The genotypic correlation coefficients were found to lie between the range -0.7675 and 0.9837. The genetic correlation between the number of capsules on main stem and number of capsules/plant, number of seeds per capsule and circumference of capsule, and number of capsules on branches and number of capsules per plant were found to be highly positive whereas the number of days for flowering and number of capsules on main stem was highly negative.

4.1.4 Factor analysis

Initially the eigen values and corresponding eigen vectors of the phenotypic, genotypic and environment correlation matrices were found out by Jacobi's method. The phenotypic and environment correlation

Table 4.1.1 Phenotypic correlation matrix - data A

	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
X_1	0.0158	0.5657	0.5608	0.6055	0.7045	0.2592	0.4082	0.1732	0.0895	-0.2853	0.0755	-0.1171	0.1509	0.5285
X_2		0.3945	-0.3703	0.3422	-0.0460	-0.1050	-0.1711	-0.2574	-0.3018	0.3399	-0.0045	-0.0296	-0.0290	-0.0047
X_3			-0.0494	0.3024	0.1522	0.1443	0.1068	-0.0581	-0.1076	0.2664	0.1389	0.0216	0.0421	0.0738
X_4				0.3899	0.8581	0.3441	0.2617	0.1623	0.1529	-0.5071	-0.0523	-0.2313	0.0647	0.6010
X_5					0.7970	0.1302	0.2225	0.0077	-0.0818	-0.1562	-0.0631	-0.1940	0.0657	0.6063
X_6						0.2985	0.2982	0.1162	0.0546	-0.4157	-0.0762	-0.2531	0.0849	0.7191
X_7							0.0390	0.0646	0.0203	-0.1294	0.0614	-0.0811	0.1948	0.1891
X_8								0.4146	0.3610	-0.2820	0.1107	-0.0354	0.1617	0.2676
X_9									0.6369	-0.0538	-0.0105	0.0364	0.0020	0.1708
X_{10}										-0.0729	0.1619	-0.0832	0.0613	0.1179
X_{11}											0.2920	-0.0101	-0.0550	-0.3269
X_{12}												-0.1385	0.1147	-0.0670
X_{13}													0.0442	-0.1852
X_{14}														0.0903

Table 4.1.3 Environment correlation matrix – data A

	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
X_1	0.1804	0.7424	0.6744	0.5529	0.7050	0.3168	0.2416	0.0514	0.0680	-0.3092	0.0251	-0.1560	-0.0836	0.6300
X_2		0.3129	0.1527	0.2945	0.2735	0.0693	0.1363	0.1487	0.0799	-0.2426	[illegible]	[illegible]	-0.0575	0.2094
X_3			0.3827	0.3842	0.4572	0.2823	0.1541	-0.0339	0.0616	-0.2694	0.0055	-0.2025	-0.0756	0.4067
X_4				0.5223	0.8793	0.3722	0.1347	0.0239	0.0256	-0.1041	-0.0227	-0.0339	-0.0342	0.7621
X_5					0.8482	0.1894	0.1480	0.0625	0.0641	-0.1902	-0.1091	-0.0869	-0.0320	0.8072
X_6						0.3478	0.1594	0.0629	0.0419	-0.1737	-0.0740	-0.0715	-0.0846	0.8868
X_7							-0.0536	0.0161	-0.0999	-0.0455	0.0364	0.0839	-0.0505	0.3064
X_8								0.4504	0.2729	-0.2180	0.1778	0.0451	0.0040	0.1599
X_9									0.3226	-0.0680	0.0967	0.2058	0.0865	0.0949
X_{10}										-0.0534	0.2106	0.1887	-0.0757	0.0907
X_{11}											0.1082	-0.0299	-0.0089	-0.1896
X_{12}												0.1981	-0.0723	0.0047
X_{13}													0.1381	-0.0169
X_{14}														-0.0255

Table 4.1.4 Genotypic correlation matrix – data R

	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
X_1	-0.0756	0.4399	0.4690	0.6802	0.7168	0.2598	0.6610	0.2804	0.1107	-0.2767	0.1349	-0.1231	0.2039	0.4601
X_2		0.4424	-0.7030	0.4110	-0.2826	-0.5020	-0.5018	-0.5273	-0.5870	0.5808	-0.0828	-0.0242	-0.0323	-0.1147
X_3			-0.3905	0.2311	-0.1298	-0.1056	0.0584	-0.0789	-0.2618	0.5532	0.2846	0.0877	0.0586	-0.1447
X_4				0.2423	0.8418	0.4330	0.4603	0.2958	0.2853	-0.7675	-0.0900	-0.3267	0.0949	0.4838
X_5					0.7277	-0.0418	0.3629	-0.0602	-0.2761	-0.1441	0.0096	-0.2876	0.1063	0.4400
X_6						0.2429	0.5426	0.1774	0.0702	-0.6188	-0.0797	-0.3761	0.1341	0.5905
X_7							0.4554	0.2370	0.4110	-0.3903	0.1727	-0.3509	0.6135	0.0359
X_8								0.3798	0.5174	-0.3954	-0.0192	-0.0912	0.2927	0.4265
X_9									0.9837	-0.0468	-0.1510	-0.0184	-0.0025	0.2339
X_{10}										-0.0906	0.0949	-0.2010	0.0957	0.1440
X_{11}											0.4769	-0.0060	-0.0650	-0.3989
X_{12}												-0.3280	0.1935	-0.1444
X_{13}													0.0440	-0.2466
X_{14}														0.1166

matrices were found to be non-negative definite but genotypic correlation matrix was in indefinite form. The latent roots of the phenotypic and environment correlation matrices are given in Table 4.1.5. The variables with high genotypic correlation coefficients were eliminated to make it in positive definite form and the latent roots of the resulting positive definite genotypic correlation matrix is presented in Table 4.1.6.

The first five latent roots of the phenotypic and environment correlation matrices were greater than one and they altogether contributed 71.03 percent and 67.59 percent respectively to the total variation. The first four eigen values of genotypic correlation matrix were found to be greater than one and explained 74.66 percent of the total variation.

4.1.4.1 Principal factor analysis

PFA of the phenotypic correlation matrix of order 15 was done with the squared multiple correlation coefficients (SMC) as first estimates of communalities and a five factor solution was extracted. Twelve iterations were needed for the convergence of communalities with a difference of five units in the third decimal place. The

Table 4.1.5 Latent roots of phenotypic and environment correlation matrix - data - A

Sl. No.	Phenotypic correlation matrix		Environment correlation matrix	
	Latent roots	Percent contribution to variance	Latent roots	Percent contribution to variance
1.	4.2690	28.4600	4.6827	31.2180
2.	2.2925	15.2833	1.9310	12.8733
3.	1.7180	11.4533	1.3102	8.7347
4.	1.2156	8.1040	1.1716	7.8107
5.	1.1606	7.7373	1.0361	6.9073
6.	0.8903	5.9353	0.8891	5.9273
7.	0.8070	5.3800	0.8472	5.6480
8.	0.6578	4.3853	0.7928	5.2853
9.	0.5397	3.5980	0.6963	4.6420
10.	0.4229	2.8193	0.5011	3.3407
11.	0.3503	2.3353	0.4334	2.8893
12.	0.2938	1.9587	0.4200	2.8000
13.	0.2452	1.6347	0.1663	1.1087
14.	0.1309	0.8727	0.1125	0.7500
15.	0.0063	0.0420	0.0098	0.0653

Table 4.1.6 Latent roots of genotypic correlation matrix of order 10 - data A

Sl. No.	Latent roots	Percent contribution to variance
1.	2.9770	29.7697
2.	1.7989	17.9895
3.	1.5026	15.0259
4.	1.1866	11.8663
5.	0.9985	9.9848
6.	0.6107	6.1066
7.	0.5087	5.0868
8.	0.2072	2.0721
9.	0.1898	1.8985
10.	0.0199	0.1998

principal factor loadings in the 12th iteration along with communalities in the 11th and 12th iterations are summarised in Table 4.1.7. The loadings in the 3rd and 4th factors in the 12th iteration lead to unsatisfactory results and hence varimax rotation of loadings was applied and the results are given in Table 4.1.8. The initial and final estimates of communalities have similarity except for the 12th and 13th variables. The important characters contributing to each factor were isolated in accordance with the procedure given in Harman (1967).

Factor I :	number of capsules per plant
	number of capsules on branches
	number of capsules on mainstem
	yield of seeds per plant
Factor II :	number of branches
Factor III :	height of the plant
	height upto first capsule
Factor IV :	circumference of capsule and
	number of seeds per capsule
	length of the capsule
Factor V :	number of days to flowering
	number of days to maturity

Table 4.1.7 Principal factor solution in the 12th iteration for the phenotypic correlation matrix - data A

Variable	Common factor coefficients					Estimated communality		Original communality (SMC)
	1	2	3	4	5	12th iteration	11th iteration	
1	0.8185	0.2364	0.2514	0.2682	0.2092	0.9047	0.9037	0.7879
2	-0.1144	0.6979	-0.2930	0.1333	0.0019	0.6038	0.6053	0.5256
3	0.2174	0.5941	0.2552	0.4760	0.1719	0.7215	0.7200	0.6688
4	0.8286	-0.2453	0.2012	-0.2789	-0.1660	0.8926	0.8921	0.9762
5	0.7204	0.4645	-0.3540	-0.0160	-0.0332	0.8614	0.8604	0.9654
6	0.9649	0.1108	-0.0776	-0.1944	-0.1323	1.005	1.004	0.9893
7	0.3086	-0.0038	0.2595	-0.0045	-0.0853	0.1699	0.1698	0.1995
8	0.4302	-0.2477	-0.0294	0.3243	0.1570	0.3771	0.3766	0.3789
9	0.2568	-0.5141	-0.2468	0.4863	0.0981	0.6372	0.6390	0.5147
10	0.1981	-0.5791	-0.1641	0.5037	-0.1176	0.6691	0.6683	0.4804
11	-0.4583	0.3433	-0.0330	0.3940	-0.3253	0.5901	0.5855	0.5061
12	-0.0298	0.0356	0.2347	0.3572	-0.3768	0.3269	0.3276	0.7609
13	-0.2151	-0.0415	0.0377	0.0192	0.3729	0.1888	0.1928	0.8565
14	0.1347	-0.0137	0.1505	0.0920	0.0041	0.0495	0.0496	0.1072
15	0.7332	0.0343	-0.1707	-0.0793	-0.0595	0.5632	0.5633	0.5449

Table 4.1.8 Rotated principal factor loadings for the phenotypic correlation matrix - data A

Variable	Common factor coefficients				
	1	2	3	4	5
1	0.5616	0.2364	0.6544	0.2891	0.1469
2	-0.0325	0.6979	-0.2574	0.2168	-0.0491
3	-0.0321	0.5941	0.4560	0.3904	0.0847
4	0.7413	-0.2453	0.4502	-0.2617	-0.1089
5	0.7736	0.4645	0.0134	0.2012	-0.0815
6	0.9402	0.1108	0.3003	-0.0558	-0.1228
7	0.1814	-0.0038	0.3537	-0.0853	-0.0676
8	0.3028	-0.2477	0.2632	0.3871	0.0701
9	0.1825	-0.5141	0.0524	0.5793	-0.0357
10	0.0954	-0.5791	0.1017	0.5066	-0.2401
11	-0.5052	0.3433	-0.1099	0.2318	-0.3888
12	-0.2113	0.0356	0.2937	0.1370	-0.4194
13	-0.2103	-0.0415	-0.0671	0.0619	0.3685
14	0.0394	-0.0138	0.2144	0.0414	-0.0055
15	0.7271	0.0343	0.1502	0.0692	-0.0774
Proportionate variance accounted by each factor	0.3260	0.2139	0.1482	0.1325	0.0649

The PFA method was applied to the environment correlation matrix of order 15 and fifty-five iterations were taken for the convergence of communalities with a difference of five units in the third decimal place. This matrix was singular and hence SMC's were not estimable by equation (7) so the largest correlation coefficient in each array was taken as the initial estimates of corresponding array communality (Harman, 1967). The PF loadings in the 55th iteration is given in Table 4.1.9 along with communalities in the 54th and 55th iterations. Varimax rotation of the loadings helped to derive meaningful interpretation of the factor loadings and the results are summarised in Table 4.1.10. The variables with high loadings in each factor are given below.

	number of cap/plant
	yield of seeds/plant
Factor I :	number of capsules on branches
	number of capsules on main stem
Factor II :	height upto first capsule
	height of the plant
	circumference of capsule
Factor III :	length of the capsule
	number of seeds per capsule

Table 4.19 Principal factor solution in the 55th iteration for the environment correlation matrix - data A

Variable	Common factor coefficients					Estimated communality		Original communality
	1	2	3	4	5	55th iteration	54th iteration	
1	0.8278	0.3395	0.0072	0.0473	0.1512	0.8257	0.8257	0.7424
2	0.3039	0.1497	0.1506	-0.1088	-0.2075	0.1923	0.1923	0.3129
3	0.6362	0.6128	-0.0302	-0.0071	-0.0107	0.7813	0.7813	0.7424
4	0.8210	-0.1808	-0.0979	0.2207	0.2981	0.8538	0.8539	0.8793
5	0.8163	-0.2371	-0.0705	-0.2045	-0.3930	0.9238	0.9240	0.8482
6	0.9656	-0.2469	-0.1022	0.0325	-0.0060	1.0049	1.0049	0.8868
7	0.3530	0.0097	-0.0522	0.1974	0.1738	0.1966	0.1966	0.3722
8	0.2364	0.0773	0.5483	-0.1982	0.0500	0.4043	0.4044	0.4504
9	0.1045	-0.1472	0.6541	-0.2709	0.1707	0.5630	0.5627	0.4504
10	0.0952	-0.0197	0.4623	-0.0054	-0.0068	0.2233	0.2236	0.3226
11	-0.2673	-0.2097	-0.1074	0.2455	0.0074	0.1873	0.1874	0.1082
12	-0.0201	0.0820	0.5188	0.6267	-0.2503	0.7317	0.7267	0.2106
13	-0.0850	-0.2186	0.3070	0.1042	0.1035	0.1708	0.1709	0.2088
14	-0.0817	-0.0930	0.0341	-0.1165	0.0615	0.0339	0.0340	0.1381
15	0.8725	-0.2534	-0.0102	0.0420	-0.0608	0.8310	0.8310	0.8868

Table 4.1.10 Rotated principal factor loadings for the environment correlation matrix - data A

Variable	Common factor coefficients				
	1	2	3	4	5
1	0.5764	0.6559	0.0362	0.1252	0.2154
2	0.1999	0.2790	0.1357	0.0890	0.2195
3	0.2799	0.8196	-0.0140	0.1642	0.0630
4	0.8113	0.1739	-0.0859	0.0880	0.3876
5	0.8333	0.2127	-0.0891	-0.0035	-0.4199
6	0.9701	0.2229	-0.1021	0.0487	0.0358
7	0.3084	0.0289	-0.0609	0.1202	0.3584
8	0.1737	0.1768	0.5821	0.0040	-0.0628
9	0.1608	-0.0685	0.7179	-0.1296	-0.0154
10	0.0935	-0.0053	0.4422	0.1318	-0.0410
11	-0.1398	-0.3467	-0.1645	0.1106	0.0907
12	-0.0559	-0.1103	0.3019	0.7907	0.0062
13	0.0259	-0.2714	0.2833	0.0858	0.0942
14	-0.0293	-0.0945	0.0693	-0.1386	-0.0067
15	0.8906	0.1666	-0.0277	0.0942	-0.0206
Proportionate variance accounted by each factor	0.4923	0.2096	0.1724	0.1036	0.0766

Factor IV : number of days to maturity

Factor V : number of branches
number of fruiting nodes per 20 cm

A three factor-model was fitted to the genotypic correlation matrix of order 10 by the PFA method. The largest correlation coefficient in each array was taken as the first estimate of the corresponding array communality. Twenty-nine iterations were taken for the convergence of communalities with a 5 unit difference in the third decimal place. The PF loadings in the 29th iteration along with communalities in the 28th and 29th iterations are given in Table 4.1.11 and the rotated loadings in 4.1.12. The variables influencing substantially each factor are shown below.

Factor I : number of fruiting nodes/20 cm
length of capsule
circumference of capsule

Factor II : height of the plant
height upto first capsule
number of branches

Table 4.1.11 Principal factor solution in the 29^{th} iteration for the genotypic correlation matrix of order 10 - data A

Variable Code	Common factor coefficients			Estimated communality		Original communality
	1	2	3	29^{th} iteration	28^{th} iteration	
1	0.7076	0.7303	-0.1510	1.0567	1.0563	0.6610
2	-0.5702	0.5145	0.2861	0.6717	0.6719	0.4424
3	-0.0178	0.6572	0.2643	0.5021	0.5021	0.4424
7	0.8683	-0.3857	0.6120	1.2773	1.2725	0.6135
8	0.7819	0.1335	-0.2452	0.6894	0.6895	0.6610
9	0.4491	-0.1318	-0.3682	0.3547	0.3547	0.3798
12	0.1151	0.1369	0.3787	0.1754	0.1759	0.2846
13	-0.2137	0.0001	-0.6184	0.0740	0.0741	0.0877
14	0.3969	0.0149	0.3588	0.2865	0.2871	0.6135
15	0.3804	0.1485	-0.3196	0.2689	0.2689	0.4601

Table 4.1.12 Rotated principal factor loadings for the genotypic correlation matrix - data A

Variable Code	Common factor coefficients		
	1	2	3
1	0.7076	0.6867	-0.2907
2	-0.5702	0.5605	0.1801
3	-0.0179	0.6962	0.1308
7	0.8683	-0.2588	0.6756
8	0.7819	0.0831	-0.2666
9	0.4491	-0.2021	-0.3353
12	0.1151	0.2082	0.3447
13	-0.2137	-0.0328	-0.1652
14	0.3969	0.0847	0.3490
15	0.3804	0.0832	-0.3424
Proportionate variance explained by each factor	0.5928	0.3406	0.1495

number of days to maturity

Factor III : ~~number of fruiting nodes per 20 cm~~

oil content

4.1.4.2 Maximum-Likelihood factor analysis

From the principal factor analysis of the data it was hypothesized that a minimum of five factors would suffice to describe the dependence structure parsimoniously in the case of 15 variable study. The ML solutions were extracted successively for each factor-model by Lawley's iterative scheme. The sequence terminates either when a proper acceptable solution has been found from the point of view of goodness of fit or when the number of factors becomes equal to a given upper bound.

The phenotypic correlation matrix of order 15 was subjected to ML factor analysis initially with five factors and then increasing the factors successively to get the appropriate factor-model. The approximate χ^2-statistics for the solutions are presented below.

No. of factors	Chi-square	Degrees of freedom	No. of iterations
5	66.24	40	20
6	50.19	30	32
7	44.55	21	209
8	11.84	13	77

The initial estimates of factor loadings and unique variances were obtained from the principal component method of factor analysis (vide Table 4.1.13). Seventy-six iterations were required for convergence with a five unit difference in the third decimal place. The ML estimates of factor loadings and unique variances in the 76th and 77th iterations are presented in Tables 4.1.14 and 4.1.15 respectively. The varimax rotation of the factors aided to interpret the factors meaningfully and are summarised in Table 4.1.16. The variables which were highly correlated with the factors are given below.

	number of cap/plant
	number of capsules on main stem
Factor I :	number of capsules on branches
	yield of seeds/plant
	length of capsule
Factor II :	circumference of capsule
	number of seeds/cap
Factor III :	height of the plant
	height upto first capsule
Factor IV :	number of branches

Table 4.1.13 Initial estimates of factor loadings and corresponding unique variances for 8 factors of the phenotypic correlation matrix - data A

Variable	Factor loadings								Unique variance
	1	2	3	4	5	6	7	8	
1	0.8121	0.2476	0.2290	0.0318	0.1661	0.1258	0.2236	0.0883	0.8754
2	-0.1356	0.7664	0.1791	0.2650	-0.0421	-0.1635	-0.2731	0.0373	0.8126
3	0.2177	0.6152	0.5024	0.0593	0.2285	0.3046	0.1606	0.2007	0.8929
4	0.8285	-0.1893	-0.2743	-0.1809	-0.0955	0.1261	0.0930	-0.1113	0.8762
5	0.7145	0.4841	-0.0150	0.2190	-0.0866	-0.1592	-0.1264	-0.1289	0.8584
6	0.9326	0.1474	-0.1773	0.0071	-0.0957	-0.0003	-0.0120	-0.1286	0.9489
7	0.3733	0.0198	0.0122	-0.5873	0.2062	0.4719	-0.3505	0.1084	0.8846
8	0.4986	-0.3133	0.4018	0.2006	0.1627	-0.2486	0.2672	0.2904	0.7924
9	0.2866	-0.5741	0.4660	0.3294	-0.0072	0.1666	-0.3047	-0.0384	0.8595
10	0.2248	-0.6341	0.5027	0.1309	-0.1661	0.0161	-0.1994	-0.0393	0.7916
11	-0.4990	0.3584	0.5047	-0.0563	-0.2275	0.1117	-0.2081	-0.2139	0.7886
12	-0.0302	0.0368	0.5791	-0.4911	-0.2734	-0.1092	0.3863	-0.3420	0.9317
13	-0.2560	-0.0839	-0.0083	0.2611	0.7662	0.1442	0.1069	-0.4721	0.9830
14	0.1732	-0.0201	0.2007	-0.4741	0.4801	-0.5844	-0.2816	0.0398	0.9484
15	0.7817	0.0771	-0.1110	0.1093	-0.1121	-0.1203	-0.1475	-0.2768	0.7666

Table 4.1.14 Maximum likelihood estimates of factor loadings and unique variances in the 76th iteration of phenotypic correlation matrix - data A

Variable	Factor loadings								Unique variance
	1	2	3	4	5	6	7	8	
1	0.6976	-0.1511	-0.4400	-0.2044	0.0969	-0.0989	-0.0934	0.2465	0.1666
2	-0.1003	0.1841	-0.0628	-0.6675	0.0687	0.0728	-0.0927	-0.1670	0.4601
3	0.1232	-0.0021	-0.7449	-0.4696	0.2185	-0.0676	-0.3355	-0.0170	0.0443
4	0.8989	0.0150	0.0405	0.4263	0.0018	-0.0426	0.0199	-0.0067	0.0062
5	0.7459	-0.0019	0.0981	-0.6479	-0.0405	-0.0125	0.0371	0.0014	0.0111
6	0.9931	0.0005	0.0140	-0.0836	-0.0037	-0.0394	0.0029	-0.0048	0.0050
7	0.3131	0.0229	-0.2247	0.1256	0.0876	-0.0186	0.0827	-0.0823	0.8135
8	0.2996	-0.4550	-0.1629	-0.0102	-0.0567	-0.0586	0.1354	0.3249	0.5460
9	0.1247	-0.8785	-0.0152	0.1394	-0.0855	-0.0848	0.0263	-0.1141	0.1649
10	0.0772	-0.6612	-0.0855	0.2220	-0.1644	0.0623	0.1251	-0.0111	0.4536
11	-0.4349	0.0266	-0.2637	-0.2951	-0.1505	0.1623	-0.0565	-0.5192	0.3318
12	-0.0645	0.0662	-0.5456	-0.0038	-0.7134	0.2111	0.2700	0.0179	0.0671
13	-0.2940	0.0115	0.0006	-0.0170	-0.0567	-0.9448	0.0039	-0.0039	0.0172
14	0.0829	0.0127	-0.2964	-0.0230	0.3515	-0.0898	0.8428	-0.0194	0.0623
15	0.7164	-0.1092	0.0526	-0.0956	-0.0070	-0.0259	0.0548	0.0355	0.4581

Table 4.15 Maximum likelihood estimates of factor loadings and unique variances in the 77th iteration of phenotypic correlation matrix - data A

Variable	Factor loadings								Unique variance
	1	2	3	4	5	6	7	8	
1	0.6975	-0.1509	-0.4405	-0.2051	0.0954	-0.0992	-0.0907	0.2466	0.1667
2	-0.1008	0.1838	-0.0633	-0.6675	0.0683	0.0725	-0.0924	-0.1672	0.4602
3	0.1229	-0.0023	-0.7474	-0.4701	0.2160	-0.0681	-0.3313	-0.0169	0.0440
4	0.8992	0.0149	-0.0407	0.4256	0.0017	-0.0424	-0.0197	-0.0067	0.0062
5	0.7454	-0.0019	0.0988	-0.6484	-0.0403	-0.0128	0.0367	0.0013	0.0110
6	0.9930	0.0005	0.0142	-0.0843	-0.0037	-0.0394	0.0029	-0.0048	0.0050
7	0.3133	0.0227	-0.2244	0.1253	0.0873	-0.0186	0.0839	-0.0824	0.8135
8	0.2996	-0.4546	-0.1618	-0.0105	-0.0571	-0.0587	0.1366	0.3255	0.5460
9	0.1248	-0.8790	-0.0148	0.1393	-0.0860	-0.0848	0.0266	-0.1134	0.1639
10	0.0774	-0.6608	-0.0845	0.2219	-0.1646	0.0624	0.1260	-0.0102	0.4540
11	-0.4351	0.0261	-0.2637	-0.2949	-0.1512	0.1621	-0.0545	-0.5191	0.3319
12	-0.0644	0.0665	-0.5423	-0.0039	-0.7140	0.2110	0.2749	0.0179	0.0671
13	-0.2940	0.0116	0.0009	-0.0162	-0.0568	-0.9448	0.0041	-0.0387	0.0172
14	0.0829	0.0125	-0.2919	-0.0230	0.3527	-0.0898	0.8439	-0.0194	0.0623
15	0.7164	-0.1081	0.0531	-0.0962	-0.0068	-0.0260	0.0546	0.0356	0.4581

Table 4.1.16 Rotated maximum likelihood estimates of factor loadings of the phenotypic correlation matrix - data A

Variable	Factor loadings							
	1	2	3	4	5	6	7	8
1	0.6975	-0.1509	-0.4873	-0.0702	-0.0997	-0.0992	0.0338	0.2466
2	-0.1008	0.1838	-0.2951	-0.6089	0.0530	0.0725	-0.0460	-0.1672
3	0.1229	-0.0023	-0.9368	-0.2140	-0.0761	-0.0681	-0.0285	-0.0169
4	0.8992	0.0150	0.0796	0.4203	-0.0122	-0.0424	-0.0099	-0.0068
5	0.7454	-0.0019	-0.0751	-0.6539	0.0020	-0.0128	-0.00002	0.0014
6	0.9930	0.0005	-0.0098	-0.0850	0.0027	-0.0394	-0.0014	-0.0048
7	0.3133	0.0227	-0.1558	0.1765	-0.0529	-0.0186	0.1499	-0.0824
8	0.2996	-0.4546	-0.0739	0.0108	-0.1598	-0.0387	0.1308	0.3255
9	0.1248	-0.8790	0.0662	0.1257	-0.0866	-0.0848	-0.0063	-0.1135
10	0.0774	-0.6608	0.0912	0.2044	-0.2119	0.0624	0.0659	-0.0102
11	-0.4351	0.0261	-0.2660	-0.2290	-0.2356	0.1621	-0.0602	-0.5191
12	-0.0644	0.0665	-0.1101	0.0284	-0.9281	0.2110	0.0716	0.0179
13	-0.2940	0.0116	0.0178	-0.0222	-0.0486	-0.9448	-0.0184	-0.0039
14	0.0829	0.0126	-0.1078	0.0078	-0.0629	-0.0898	0.9522	-0.0194
15	0.7164	-0.1081	0.0367	-0.1111	0.0043	-0.0260	0.0373	0.0336
Contribution of each factor	3.8715	1.4910	1.3530	1.1880	1.0230	1.0110	0.9705	0.4860
Proportionate variance accounted by each factor	0.2581	0.0994	0.0902	0.0792	0.0682	0.0674	0.0647	0.0324

Factor V : number of days to maturity

Factor VI : 1000 seed weight

Factor VII : seed oil content

Factor VIII : number of days to flowering

The residual matrix after removal of 8 factors is given in Table 4.1.17. The largest element of the residual matrix was 0.0398. The goodness of fit test is accepted for the 8 factor-model.

The ML solutions were estimated for the environment correlation matrix of order 15, starting from a five factor model. The goodness of fit of the models with 5, 6, 7 and 8 factors are given below.

Number of factors	Chi-square	Degrees of freedom	Number of iterations
5	89.43	40	17
6	62.26	30	18
7	54.58	21	22
8	31.13	13	37

Table 4.17 Residual matrix after removal of eight factors from phenotypic correlation matrix - data A

0														
-0.0175	0													
0.0007	0.0029	0												
0.0008	-0.0009	-0.0001	0											
0.0014	-0.0016	-0.0002	-0.0000	0										
-0.0011	0.0011	0.0001	-0.0000	-0.0000	0									
-0.0113	-0.0188	0.0033	-0.0002	0.0005	0.0001	0								
-0.0110	0.0006	0.0022	-0.0005	-0.0010	0.0010	-0.0603	0							
0.0058	0.0043	-0.0013	-0.0004	-0.0005	0.0004	0.0191	0.0001	0						
-0.0200	-0.0130	0.0039	0.0006	0.0010	-0.0006	-0.0315	0.0063	0.0009	0					
-0.0009	-0.0155	0.0011	0.0011	0.0017	-0.0013	-0.0378	-0.0083	0.0040	-0.0123	0				
0.0004	0.0017	-0.0002	0.0000	-0.0001	0.0001	0.0035	0.0007	-0.0003	0.0007	0.0002	0			
-0.0001	-0.0001	0.0000	0.0000	0.0000	0.0000	-0.0003	0.0000	0.0000	-0.0001	-0.0001	0.0000	0		
0.0004	0.0003	-0.0001	0.0000	0.0000	0.0000	0.0010	0.0009	-0.0004	0.0007	0.0008	-0.0001	0.0000	0	
0.0105	0.0399	-0.0016	0.0018	0.0019	-0.0021	-0.0104	-0.0095	0.0004	0.0108	-0.0021	-0.0003	0.0001	0.0002	0

Chi-square value revealed that an eight factor-model with a 0.005 convergence criterion can be accepted at the 0.001 probability level. Thirty-seven iterations were taken for the convergence. The initial estimates of factor loadings and unique variances obtained from the principal component method of factor analysis are shown in Table 4.1.18. The ML solutions in the 36th and 37th iterations are summarised in Tables 4.1.19 and 4.1.20 respectively. The varimax rotated loadings are presented in Table 4.1.21. The residual correlation matrix after removal of 8 factors is given in Table 4.1.22. The variables influencing the eight factors are

Factor I : number of capsules per plant
number of capsules on main stem
yield of seeds/plant

Factor II : height of the plant
height upto first capsule

Factor III: length of the capsule
circumference of capsule
number of seeds per capsule

Factor IV: number of branches

Table 4.1.18 Initial estimates of factor loadings and corresponding unique variances for 8 factors of the environment correlation matrix - data A

Variable	Factor loadings								Unique variance
	1	2	3	4	5	6	7	8	
1	0.8452	-0.0310	0.1245	0.0979	0.1322	0.2854	0.0529	-0.1571	0.8668
2	0.3644	0.2014	0.4157	0.0555	0.1737	-0.6696	0.1908	0.2865	0.9463
3	0.6609	-0.0657	0.3688	0.1935	0.3022	0.2032	0.1618	-0.1246	0.7889
4	0.8268	-0.1186	-0.2762	0.0389	-0.0875	0.0724	-0.0295	0.0049	0.7892
5	0.8090	-0.0619	-0.0532	-0.1824	-0.2536	-0.2409	0.0241	-0.0449	0.8194
6	0.9412	-0.1037	-0.1816	-0.0666	-0.1689	-0.0922	-0.0085	0.0092	0.9712
7	0.4135	-0.1586	-0.3585	0.1781	0.5337	0.1590	-0.1336	0.4133	0.8552
8	0.2719	0.6610	0.2238	-0.0653	-0.1244	0.3511	0.0908	0.1428	0.7326
9	0.1193	0.7249	-0.0578	-0.2174	-0.1077	0.1334	-0.0121	0.4597	0.8312
10	0.1069	0.6541	-0.0148	0.1848	-0.2439	-0.0191	-0.1619	-0.3216	0.6631
11	-0.3227	-0.1729	-0.5074	0.2903	-0.3855	0.0352	0.4278	0.2174	0.8559
12	-0.0170	0.4488	-0.2116	0.5896	0.2131	-0.1090	0.3943	-0.1997	0.8468
13	-0.0993	0.4503	-0.5527	-0.1369	0.3580	-0.2277	-0.3130	-0.1860	0.8494
14	-0.1018	0.0791	-0.1766	-0.7107	0.2771	0.0604	0.5397	-0.1869	0.9595
15	0.8830	-0.0272	-0.2294	-0.0958	-0.1623	-0.1038	0.0109	-0.0754	0.8852

Table 4.1.19 Maximum likelihood estimates of factor loadings in the 36th iteration of the environment correlation matrix - data A

Variable	Factor loadings								Unique variance
	1	2	3	4	5	6	7	8	
1	0.7264	-0.6030	0.0147	-0.0498	0.0153	0.0478	-0.0049	-0.1038	0.0927
2	0.2789	0.0694	-0.0476	-0.8768	0.1618	0.0040	-0.0016	0.2498	0.0578
3	0.4712	-0.6530	-0.1192	-0.2311	0.0732	-0.0428	0.0340	0.0277	0.2748
4	0.8894	0.0330	-0.0029	-0.0226	-0.5041	0.0182	-0.0078	-0.4396	0.01111
5	0.8448	0.0090	-0.0032	0.0854	0.0248	0.0094	-0.0109	0.5147	0.0131
6	0.9949	0.0225	0.0008	0.0091	-0.0039	-0.0140	0.0151	0.0137	0.0089
7	0.3461	-0.1313	-0.0455	-0.0222	-0.0039	-0.0721	0.0142	-0.1823	0.3341
8	0.1679	-0.1765	0.5151	-0.1368	-0.0082	0.1630	-0.0461	0.0349	0.6267
9	0.0586	0.0051	0.8322	-0.1797	-0.0842	-0.0597	0.1546	0.0705	0.2322
10	0.0502	-0.0514	0.4199	-0.0748	0.0809	0.1225	-0.2099	0.0445	0.7453
11	-0.1809	0.2683	-0.0402	0.2029	-0.0128	-0.0649	-0.1079	-0.1151	0.8232
12	-0.0651	-0.0788	0.2197	-0.1484	0.0823	-0.0748	-0.6235	-0.0880	0.5104
13	-0.0731	0.1447	0.2798	-0.0009	-0.1507	-0.1776	-0.2344	-0.0456	0.7842
14	-0.0896	-0.0115	-0.0195	-0.0974	-0.9586	0.0009	-0.0038	0.1473	0.0413
15	0.8943	0.0029	0.0628	0.0567	-0.0507	-0.0225	-0.1075	0.0863	0.1711

Table 4.1.20 Maximum likelihood estimates of factor loadings and unique variances in the 37th iteration of environment correlation matrix - data A

Variable	Factor loadings								Unique variance
	1	2	3	4	5	6	7	8	
1	0.7265	-0.6033	0.0142	-0.0496	0.0159	0.0474	-0.0049	-0.1033	0.0923
2	0.2787	0.0696	-0.0477	-0.8777	0.1602	0.0041	-0.0017	0.2479	0.0578
3	0.4710	-0.6527	-0.1197	-0.2313	0.0732	-0.0433	0.0339	0.0278	0.2751
4	0.8897	0.0328	-0.0028	-0.0213	-0.0483	0.0179	-0.0077	-0.4397	0.0109
5	0.8447	0.0093	-0.0032	0.0837	0.0226	0.0093	-0.0107	0.5155	0.0129
6	0.9948	0.0225	0.0009	0.0088	-0.0038	-0.0143	0.0153	0.0140	0.0091
7	0.3460	-0.1311	-0.0452	-0.0217	-0.0029	-0.7014	0.0136	-0.1820	0.3353
8	0.1679	-0.1769	0.5147	-0.1369	-0.0084	0.1634	-0.0466	0.0345	0.6267
9	0.0585	0.0044	0.8326	-0.1800	-0.0847	-0.0595	0.1538	0.0696	0.2317
10	0.0503	-0.0515	0.4196	-0.0750	0.0807	0.1230	-0.2108	0.0446	0.7450
11	-0.1808	0.2684	-0.0399	0.2033	-0.0123	-0.0648	-0.1074	-0.1147	0.8233
12	-0.0650	-0.0788	0.2188	-0.1481	0.0826	-0.0741	-0.6227	-0.0880	0.5120
13	-0.0730	0.1445	0.2797	-0.0008	-0.1505	-0.1773	-0.2354	-0.0463	0.7839
14	-0.0895	-0.0116	-0.0196	-0.0974	-0.9594	0.0009	-0.0038	0.1425	0.0413
15	0.8943	0.0031	0.0627	0.0562	-0.0509	-0.0227	-0.1074	0.0865	0.1710

Table 4.1.21 Rotated maximum likelihood estimates of factor loadings for the environment correlation matrix - data A

Variable	Factor loadings							
	1	2	3	4	5	6	7	8
1	0.5445	-0.7716	0.0142	-0.0202	0.0159	0.0228	-0.0049	-0.1202
2	0.2872	-0.0053	-0.0477	-0.9120	0.1602	0.0050	-0.0017	0.0038
3	0.2850	-0.7528	-0.1197	-0.2304	0.0732	-0.0497	0.0339	-0.0252
4	0.8676	-0.1998	-0.0028	0.0968	-0.0483	-0.0721	-0.0077	-0.4237
5	0.8180	-0.2108	-0.0032	-0.0569	0.0226	0.1175	-0.0107	0.5058
6	0.9664	-0.2372	0.0009	0.0048	-0.0038	-0.0107	0.0153	0.0185
7	0.2999	-0.2166	-0.0452	0.0277	-0.0029	-0.7238	0.0136	-0.0308
8	0.1161	-0.2145	0.5147	-0.1412	-0.0084	0.1591	-0.0466	-0.0373
9	0.0576	-0.0109	0.8326	-0.1921	-0.0847	-0.0542	0.1538	0.0310
10	0.0351	-0.0628	0.4196	-0.0842	0.0807	0.1251	-0.2109	-0.0032
11	-0.1048	0.3062	-0.0399	0.2266	-0.0123	-0.0751	-0.1074	-0.0415
12	-0.0833	-0.0592	0.2188	-0.1192	0.0826	-0.0984	-0.6227	-0.1061
13	-0.0329	0.1585	0.2797	0.0116	-0.1505	-0.1827	-0.2354	-0.0069
14	-0.0894	0.0121	-0.0196	-0.1318	-0.9594	0.0241	-0.0038	0.1087
15	0.8643	-0.2298	0.0627	0.0311	-0.0509	-0.0016	-0.1074	0.1010
Contribution of each factor	3.6975	1.5750	1.2855	1.0470	1.0005	0.6390	0.5385	0.4890
Proportionate variance accounted by each factor	0.2465	0.1050	0.0857	0.0698	0.0667	0.0426	0.0359	0.0326

Table 4.1.11 Residual matrix after removal of eight factors from environment correlation matrix - data A

0														
-0.0001	0													
0.0005	-0.0001	0												
0.0013	0.0003	-0.0034	0											
0.0014	0.0003	-0.0038	-0.0002	0										
-0.0015	-0.0003	0.0042	0.0001	0.0001	0									
0.0005	0.0002	-0.0023	0.0006	0.0006	-0.0009	0								
-0.0055	-0.0018	-0.0022	0.0011	0.0013	-0.0004	0.0069	0							
0.0029	0.0006	-0.0041	-0.0009	-0.0010	0.0009	-0.0012	0.0019	0						
-0.0128	-0.0012	0.0425	0.0019	0.0015	-0.0020	-0.0093	-0.0017	0.0005	0					
-0.0144	-0.0038	0.0381	0.0015	0.0017	0.0007	-0.0101	-0.0823	0.0309	-0.0070	0				
0.0044	0.0006	-0.0051	-0.0029	-0.0028	0.0036	-0.0040	0.0287	-0.0035	-0.0180	0.0757	0			
0.0051	0.0001	-0.0279	0.0008	0.0008	-0.0012	0.0107	-0.0428	-0.0044	0.0688	-0.1145	-0.0079	0		
-0.0001	0.0000	0.0006	0.0000	0.0000	0.0000	-0.0002	0.0004	0.0001	-0.0007	0.0012	0.0003	0.0001	0	
-0.0057	-0.0006	0.0119	0.0029	0.0028	-0.0036	0.0026	-0.0190	0.0053	0.0042	-0.0413	0.0009	-0.0015	-0.0002	0

Factor V : seed oil content

Factor VI : number of fruiting nodes/20 cm

Factor VII : number of days to maturity

Factor VIII : number of capsules on branches

Three and four factor solutions were extracted for the genotypic correlation matrix of order 10. The goodness of fit of the factor-model for these factors are summarised below.

Number of factors	Chi-square	Degrees of freedom	Number of iterations
3	226.69	18	46
4	215.51	11	53

The tests resoundingly rejected the three and four factor models and suggested that a higher-dimensional or non-linear mechanism had generated the observed correlations.

There was no scope for the acceptance of higher-dimensional factor models. The initial estimates of loadings and unique variances for a three factor model are given in Table 4.1.23. The ML solutions in the 45th and 46th iterations are given in Tables 4.1.24 and 4.1.25 respectively. Forty-six iterations were taken for the convergence with a five unit difference in third decimal place. The rotated loadings are given in Table 4.1.26. The variables influencing the three factors are

	number of fruiting nodes/20 cm
Factor I :	circumference of capsule
	length of the capsule
	height of the plant
Factor II :	height upto first capsule
	yield of seeds per plant
Factor III :	oil content
	number of branches

The residual correlation matrix is given in Table 4.1.27

4.2 Results of data B (rice fallows)

4.2.1 Preliminary statistical analysis

The analysis of variance for simple lattice design

Table 4.1.23 Initial estimates of factor loadings and corresponding unique variances for three factors of the genotypic correlation matrix - data A

Variable Code	Factor loadings			Unique variance
	1	2	3	
1	0.6662	0.4666	0.4643	0.1228
2	-0.6252	0.5654	0.2306	0.2362
3	-0.0452	0.8129	0.3079	0.2423
7	0.7271	0.0349	-0.5477	0.1701
8	0.8469	0.0243	0.2417	0.2237
9	0.5652	-0.3841	0.2499	0.4706
12	0.1326	0.5576	-0.5038	0.4177
13	-0.2843	-0.2056	0.3277	0.7695
14	0.4646	0.3092	-0.3774	0.5460
15	0.5104	-0.0501	0.4631	0.5225

Table 4.1.24 Maximum likelihood estimates of factor loadings and unique variances in the 45th iteration of genotypic correlation matrix - data A

Variable Code	Factor loadings 1	2	3	Unique variance
1	0.4605	0.8464	0.1434	0.0510
2	-0.6877	0.3816	-0.5714	0.0548
3	-0.1080	0.6065	-0.2387	0.5635
7	0.9187	-0.1391	-0.3216	0.0331
8	0.6653	0.3662	0.3163	0.3232
9	0.4136	0.0302	0.4361	0.6378
12	0.1755	0.0650	-0.0953	0.9559
13	-0.2685	-0.0495	0.2886	0.8422
14	0.5025	0.0619	-0.4802	0.5130
15	0.1958	0.3888	0.2507	0.7477

Table 4.1.25 Maximum likelihood estimates of factor loadings and unique variances in the 46th iteration of genotypic correlation matrix - data A

Variable Code	Factor loadings			Unique variance
	1	2	3	
1	0.4572	0.8479	0.1470	0.0504
2	-0.6895	0.3815	-0.5705	0.0537
3	-0.1103	0.6069	-0.2359	0.5638
7	0.9195	-0.1341	-0.3225	0.0325
8	0.6637	0.3674	0.3176	0.3236
9	0.4135	0.0300	0.4359	0.6381
12	0.1754	0.0660	-0.0938	0.9561
13	-0.2683	-0.0519	0.2890	0.8418
14	0.5021	0.0658	-0.4800	0.5131
15	0.1940	0.3886	0.2519	0.7479

Table 4.1.26 Rotated maximum likelihood estimates of factor loadings of the genotypic correlation matrix - data A

Variable Code	Factor loadings		
	1	2	3
1	0.1921	0.9201	-0.2572
2	-0.7323	-0.1058	-0.6315
3	-0.2609	0.3918	-0.4632
7	0.9383	0.0033	-0.2951
8	0.5138	0.6362	0.0871
9	0.3611	0.3262	0.3536
12	0.1553	0.0658	-0.1243
13	-0.2603	0.0035	0.3007
14	0.4921	-0.0075	-0.4946
15	0.0619	0.4958	0.0498
Contribution of each factor	2.2539	1.7727	1.2519
Proportionate variance accounted by each factor	0.2254	0.1773	0.1252

Table 4.1.27 Residual matrix after removal of three factors - data A

0									
0.0000	0								
0.0104	0.0003	0							
0.0006	-0.0008	0.0012	0						
-0.0007	-0.0032	-0.0165	-0.0031	0					
0.0018	-0.0050	0.0513	0.0014	-0.0441	0				
0.0126	-0.0406	0.2418	-0.0099	-0.1300	-0.1846	0			
0.0010	-0.0245	0.1577	-0.0179	0.0141	-0.0319	-0.2504	0		
-0.0109	0.0150	-0.0392	0.0058	0.0877	-0.0029	0.0561	0.3208	0	
0.0049	0.0145	-0.2997	-0.0091	0.0749	0.0322	-0.1804	-0.2462	0.1145	0

was made for each character under study. Table 4.2.1 shows the F-values for testing equality of varietal means. All characters were found to distinguish the genotypes. The mean values of the types in respect of each character are presented in Appendix IV.

4.2.2 Analysis of dispersion

Using the multivariate analysis of variance the total dispersion matrix was split up into 'between' and 'within' dispersion matrices. The between and within dispersion matrices are given in Appendices V and VI respectively. The Wilk's lambda statistic was calculated as

$$\wedge = 3.5 \times 10^{-8}$$

So that V = 2430.62 is a chi-square with 1485 degrees of freedom and is significant at one percent probability level.

4.2.3 Estimated correlation matrices

Analysis of covariance was done for each pair of characters under study. The phenotypic, environment and genotypic correlation matrices were calculated and are presented in Table 4.2.2, 4.2.3 and 4.2.4 respectively. The estimated genotypic correlation coefficients lies

Table 4.2.1 F-values obtained from the analysis of variance for 15 characters - data B

Sl. No.	Character	F-values
1.	Height of the plant	2.2946**
2.	No. of branches	2.6744**
3.	Height upto first capsule	4.8476**
4.	No. of capsules on main stem	2.5646**
5.	No. of capsules on branches	2.2520**
6.	No. of capsules/plant	2.0308**
7.	No. of fruiting nods per 20 cm	3.5034**
8.	Length of capsule	7.6995**
9.	Circumference of capsule	5.5623**
10.	No. of seeds/capsule	7.6366**
11.	No. of days for flowering	6.2338**
12.	No. of days to maturity	4.9308**
13.	1000 seed weight	5.7189**
14.	Oil content	1794.758**
15.	Yield of seeds/plant	1.3797*

Table 4.1-1 Phenotypic correlation matrix - data B

	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
X_1	0.3874	0.5822	0.3415	0.4901	0.5302	0.1720	0.0278	-0.0907	-0.0549	0.2736	0.3386	0.1021	0.0525	0.3784
X_2		0.4028	-0.0738	0.6390	0.4516	-0.0808	-0.1451	-0.1619	-0.1512	0.2900	0.1797	0.0550	-0.0033	0.3092
X_3			-0.3130	0.1705	-0.0182	0.1177	-0.0130	-0.0190	-0.0786	0.6074	0.4787	0.1998	0.0204	-0.0181
X_4				0.2873	0.6861	0.2195	-0.0139	-0.1393	-0.1242	-0.3274	-0.2058	-0.0390	-0.0625	0.3821
X_5					0.8938	0.0760	-0.0026	-0.2015	-0.1822	0.0398	0.1202	0.0561	-0.0083	0.6112
X_6						0.1581	-0.0108	-0.2191	-0.1987	-0.1217	-0.0046	-0.0034	-0.0365	0.6430
X_7							0.1870	0.0678	0.0951	0.1439	0.1167	0.0217	0.0750	0.0909
X_8								0.1664	0.2780	-0.0900	-0.0569	0.0069	0.1058	0.0456
X_9									0.7803	0.0646	-0.0513	0.1052	0.2705	-0.1797
X_{10}										0.0284	-0.0634	0.0336	0.3016	-0.1017
X_{11}											0.5347	0.0265	0.0388	-0.1107
X_{12}												0.2712	-0.0669	0.0114
X_{13}													0.1808	0.0381
X_{14}														-0.0330

Table 4.1.3 Environmental correlation matrix – data B

	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
X_1	0.3091	0.5091	0.6260	0.5463	0.6667	0.3200	0.3199	0.1924	0.1361	0.0473	0.1641	0.1369	0.0869	0.4692
X_2		0.3031	0.1387	0.5151	0.4370	0.1406	-0.0186	0.0617	-0.0776	0.2017	0.1664	-0.0924	-0.0275	0.2859
X_3			0.0957	0.2337	0.2088	0.2714	0.0638	0.0207	-0.0965	0.0834	0.1642	0.1234	0.1492	0.1219
X_4				0.4437	0.7550	0.3935	0.3050	0.1840	0.0816	-0.1144	-0.0154	0.0757	-0.0265	0.4511
X_5					0.9223	0.1882	0.2043	0.2552	0.1059	0.0906	0.2768	-0.0464	-0.0089	0.6207
X_6						0.3064	0.2789	0.2641	0.1123	0.0209	0.1945	-0.0025	-0.0184	0.6509
X_7							0.1387	-0.0798	-0.0524	0.1942	0.1624	0.0584	-0.1151	0.1371
X_8								0.1872	0.3883	0.0242	0.1010	0.1233	0.0312	0.3026
X_9									0.3626	-0.2160	-0.1487	-0.1369	-0.0077	0.1045
X_{10}										-0.0427	0.0233	-0.1252	-0.0855	0.1377
X_{11}											0.2219	0.0601	-0.0375	0.0806
X_{12}												0.2100	0.0624	0.0224
X_{13}													0.0029	0.0128
X_{14}														0.0054

Table 4.2.4 Genotypic correlation matrix – data B

	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
X_1	0.4994	0.6954	-0.0435	0.4048	0.3020	0.0162	-0.1613	-0.3271	-0.1879	0.5537	0.5223	0.0680	0.0810	0.1658
X_2		0.4990	-0.3272	0.8060	0.4771	-0.2945	-0.2337	-0.3292	-0.2097	0.3708	0.1982	0.1969	-0.0039	0.4258
X_3			-0.6335	0.1286	-0.2303	0.0254	-0.0413	-0.0367	-0.0731	0.8331	0.6347	0.2476	0.0214	-0.2488
X_4				0.0856	0.5967	0.0601	-0.1946	-0.3671	-0.2500	-0.4881	-0.3585	-0.2622	-0.0909	0.2907
X_5					0.8506	-0.0390	-0.1334	-0.5676	-0.3861	0.0063	-0.0055	0.1643	-0.0126	0.6761
X_6						-0.0039	-0.2129	-0.6500	-0.4462	-0.2531	-0.1911	-0.0046	-0.0597	0.6907
X_7							0.2163	0.1497	0.1646	0.1204	0.0896	-0.0056	0.1026	0.0288
X_8								0.1603	0.2484	-0.1267	-0.1161	-0.0454	0.1189	-0.2312
X_9									0.9198	0.1727	-0.0079	0.2388	0.3204	-0.6776
X_{10}										0.0508	-0.0957	0.1051	0.3406	-0.4440
X_{11}											0.6715	0.0094	0.0462	-0.4334
X_{12}												0.3105	-0.0835	-0.0015
X_{13}													0.2388	0.1500
X_{14}														-0.0824

between the range -0.6776 and 0.9198. The genotypic correlation between the number of capsules on branches and number of capsules per plant, circumference of capsule and number of seeds per capsule, height upto first capsule and number of days to flowering, and number of days to flowering and number of days to maturity were highly positive whereas the height upto first capsule and number of days to maturity was highly negative.

4.2.4 <u>Factor analysis</u>

The phenotypic and environment correlation matrices were positive definite while the genotypic correlation matrix was in indefinite form. The eigen values and corresponding eigen vectors of the phenotypic, genotypic and environment correlation matrices were calculated by Jacobi's method. The eigen values of the phenotypic and environment correlation matrices along with contribution of each latent root to the total variation are given in Table 4.2.5. To make the genotypic correlation matrix in positive definite form, variables with high genotypic correlation coefficients were eliminated. The latent roots of the resulting genotypic correlation matrix is given in Table 4.2.6.

The first five latent roots of the phenotypic and

Table 4.2.6 Latent roots of phenotypic and environment correlation matrix - data B

Sl. No.	Phenotypic correlation matrix		Environment correlation matrix	
	Latent roots	Percent contribution to variance	Latent roots	Percent contribution to variance
1.	3.7112	24.7413	4.2418	28.2787
2.	2.6698	17.7987	1.8699	12.4660
3.	2.0150	13.4333	1.3498	8.9987
4.	1.2189	8.1260	1.2395	8.2633
5.	1.0377	6.9180	1.1863	7.9087
6.	0.8914	5.9427	0.9891	6.5940
7.	0.8122	5.4147	0.8674	5.7827
8.	0.6860	4.5733	0.7623	5.0820
9.	0.5630	3.7533	0.6249	4.1660
10.	0.4682	3.1213	0.5195	3.4633
11.	0.3421	2.2807	0.4527	3.0180
12.	0.2614	1.7427	0.4064	2.7093
13.	0.1972	1.3147	0.2797	1.8647
14.	0.1260	0.8400	0.2103	1.4020
15.	0.0001	0.0007	0.0003	0.0020

Table 4.2.6 Latent roots of the genotypic correlation matrix of order 9 - data B

Sl. No.	Latent roots	Percent contribution to variance
1.	2.2696	25.2186
2.	1.6579	18.4211
3.	1.2040	13.3778
4.	1.1238	12.4867
5.	0.8753	9.9726
6.	0.8468	9.4089
7.	0.5574	6.1933
8.	0.3326	3.6956
9.	0.1324	1.1471

environment correlation matrices were found to be greater than one and these altogether explained 71.02 percent and 65.91 percent respectively to the total variation. The first four eigen values of the genotypic correlation matrix explained 69.51 percent of the total variation.

4.2.4.1 Principal factor analysis

Using the principal factor analysis to the phenotypic correlation matrix a five factor model was fitted with largest correlation coefficient in each row as the estimate of communality. Eighty-one iterations were taken for the convergence of communalities with a five unit difference in the third decimal place. The estimates of loadings and communalities in the 81st iteration are given in Table 4.2.7. Factors in the 81st iteration was subjected to varimax rotation to have a more meaningful interpretation of the factors. The rotated loadings were presented in Table 4.2.8. The five factors can be explained as follows.

	number of capsules per plant
Factor I :	number of capsules on branches
	yield of seeds per plant

Table 4.2.7 Principal factor solution in the 81st iteration for the phenotypic correlation matrix - data B

Variable	Common factor coefficients					Estimated communality		Original communality
	1	2	3	4	5	81st iteration	80th iteration	
1	0.6776	0.3047	0.1734	0.0877	0.2924	0.6753	0.6753	0.7289
2	0.6018	0.2972	-0.0651	0.1111	-0.4193	0.6428	0.6429	0.5732
3	0.3137	0.7884	-0.0840	0.1063	0.1399	0.7579	0.7579	0.7392
4	0.4633	-0.5900	0.2193	-0.0393	0.4385	0.8048	0.8048	0.9991
5	0.8857	-0.0456	0.1428	0.0290	-0.3297	0.9165	0.9165	0.9996
6	0.9223	-0.3377	0.2281	0.0011	-0.0225	1.0172	1.0172	0.9998
7	0.1332	0.0581	0.2151	0.0485	0.3284	0.1776	0.1776	0.2107
8	-0.0679	-0.0170	0.2884	0.0314	0.0772	0.0950	0.0950	0.1637
9	-0.3562	0.1857	0.6685	0.0877	-0.0744	0.6214	0.6214	0.6396
10	-0.3679	0.1591	0.8685	0.1917	-0.1118	0.9642	0.9642	0.6702
11	0.1237	0.7094	-0.0809	0.2289	0.1085	0.5893	0.5894	0.5426
12	0.2125	0.5725	-0.0862	-0.0480	0.1734	0.4128	0.4129	0.4343
13	0.0779	0.3701	0.2099	-0.9416	-0.0214	1.0742	1.0692	0.1926
14	-0.0743	0.1101	0.3286	-0.0670	-0.0547	0.1331	0.1332	0.1601
15	0.6062	-0.2008	0.1559	-0.0337	-0.0641	0.4374	0.4374	0.4477

Table 4.2.8 Rotated principal factor loadings for the phenotypic correlation matrix - data B

Variable	Common factor coefficients				
	1	2	3	4	5
1	0.4659	0.4970	0.0008	-0.0245	0.4591
2	0.5552	0.4735	-0.2064	0.0049	-0.2602
3	-0.0040	0.8480	-0.1007	-0.0863	0.1453
4	0.4853	-0.4265	0.0278	0.0579	0.6190
5	0.9237	0.2166	-0.1175	-0.0202	-0.0462
6	0.9621	-0.0514	-0.0715	0.0130	0.2894
7	0.0590	0.1030	0.1729	0.0260	0.3646
8	0.0026	-0.0281	0.2932	0.0388	0.0820
9	-0.1482	0.0911	0.7553	0.0690	-0.1265
10	-0.0836	0.0864	0.9474	0.1768	-0.1448
11	-0.1375	0.7477	-0.0510	0.0625	0.0693
12	-0.0455	0.5833	-0.0929	-0.1837	0.1676
13	0.0348	0.1555	0.2106	-1.0021	-0.0112
14	0.0153	0.0664	0.3445	-0.0840	-0.0526
15	0.6437	-0.0217	-0.0395	-0.0296	0.1420
Proportionate variance accounted by each factor	0.3539	0.2797	0.2154	0.1293	0.1179

	height upto first capsule
	number of days to flowering
Factor II :	number of days to maturity
	height of the plant
	number of branches
	number of seeds per capsule
Factor III :	circumference of capsule
	length of capsule
Factor IV:	1000 seed weight
Factor V:	number of capsules on main stem
	number of fruiting nodes per 20 cm

Factorization of the environment correlation matrix was done by the PFA method. Thirteen iterations were taken for the convergence of successive communalities at the 0.005 level of convergence. The factor loadings and communalities in the 12th iteration along with communalities in the 13th iteration are given in Table 4.2.9. The loadings obtained after varimax rotation are provided in Table 4.2.10. The interpretation of five factors is as follows.

Table 4.2.9 Principal factor solution in the 13th iteration for the environment correlation matrix - data B

Variable	Common factor coefficients					Estimated communality		Original communality
	1	2	3	4	5	13th iteration	12th iteration	
1	0.7872	0.3323	0.0710	0.1847	0.0791	0.7755	0.7776	0.6647
2	0.4413	-0.3046	0.2989	0.1189	0.0891	0.3989	0.3995	0.5151
3	0.3708	0.2970	0.4716	0.3785	0.2953	0.6785	0.6735	0.5091
4	0.7412	0.3264	-0.2124	-0.0770	-0.4554	0.9144	0.9122	0.7550
5	0.8617	-0.4321	0.0578	0.0080	0.0731	0.9379	0.9374	0.9223
6	0.9825	-0.2044	-0.0573	-0.0439	-0.1756	1.0432	1.0437	0.9223
7	0.3520	0.2252	0.2291	-0.1684	-0.0808	0.2620	0.2620	0.3935
8	0.3710	0.2686	-0.2891	-0.2616	0.2805	0.4405	0.4420	0.3883
9	0.2554	-0.0696	-0.4768	0.2940	0.1708	0.4130	0.4134	0.3626
10	0.1732	0.0396	-0.5569	-0.1350	0.4466	0.5594	0.5569	0.3883
11	0.0566	-0.1158	0.3510	-0.3401	0.1673	0.2835	0.2836	0.2219
12	0.2077	-0.0053	0.3199	-0.2639	0.2398	0.2726	0.2726	0.2768
13	0.0492	0.2728	0.1893	-0.1735	0.0203	0.1432	0.1433	0.2100
14	0.0088	0.0777	0.0762	0.1445	0.0605	0.0378	0.0379	0.1492
15	0.645	-0.1250	-0.0778	-0.0914	-0.0333	0.4466	0.4467	0.6509

Table 4.2.10 Rotated principal factor loadings for the environment correlation matrix - data B

Variable	Common factor coefficients				
	1	2	3	4	5
1	0.5092	0.6533	-0.2813	-0.0667	0.0759
2	0.4348	0.1426	0.1699	0.3431	-0.0676
3	0.1685	0.6929	-0.0259	0.4114	-0.0137
4	0.4726	0.4231	-0.2557	-0.6627	-0.0868
5	0.9617	0.0873	0.0041	0.0519	0.0522
6	0.9501	0.2310	-0.0959	-0.2788	-0.0116
7	0.0888	0.3919	0.1518	-0.2944	0.1095
8	0.1831	0.1631	-0.2975	-0.1941	0.5042
9	0.2554	-0.1139	-0.5734	0.0777	0.0024
10	0.1291	-0.1913	-0.5020	-0.0029	0.5042
11	0.1075	0.0486	0.4284	0.0874	0.2799
12	0.1816	0.1953	0.2945	0.1053	0.3221
13	-0.0962	0.2782	0.1404	-0.1015	0.1631
14	-0.0319	0.1292	-0.0492	0.1258	-0.0424
15	0.6187	0.1272	-0.0833	-0.1761	0.0984
Proportionate variance accounted by each factor	0.4213	0.2112	0.1602	0.1242	0.1039

Factor I :	number of capsules on main stem number of capsules per plant number of capsules on branches yield of seeds per plant
Factor II :	height upto first capsule height of the plant
Factor III :	circumference of capsule length of the capsule number of seeds per capsule
Factor IV :	number of branches number of fruiting nodes per 20 cm
Factor V :	number of days for flowering number of days to maturity

PFA method was applied to the genotypic correlation matrix of order 9. A three-factor model was fitted and 103 iterations were taken for the convergence of communalities at the derived level of five unit difference in the third decimal place. The loadings in the 103rd iteration are given in Table 4.2.11. The rotated loadings are presented in Table 4.2.12. The factors were interpreted as

Table 8.2.11 Principal factor solution in the 103^{rd} iteration for the genotypic correlation matrix of order 9 - data B

Variable Code	Common factor coefficients			Estimated communality		Original communality
	1	2	3	103^{rd} iteration	102^{nd} iteration	
1	0.5014	0.0952	0.1789	0.2924	0.2928	0.4994
2	0.8333	0.8434	-0.1619	1.4318	1.4319	0.4994
4	-0.4610	0.2253	0.8325	0.9564	0.9572	0.2907
7	-0.0872	-0.2272	0.0609	0.0629	0.0630	0.2163
8	-0.1696	-0.2228	-0.2871	0.1608	0.1608	0.2163
12	1.0427	-0.7311	0.3444	1.7404	1.7354	0.5223
13	0.2954	-0.0358	-0.1069	0.0999	0.1000	0.3105
14	0.0079	-0.0293	-0.1580	0.0259	0.0258	0.2388
15	0.1808	0.3811	0.3282	0.2857	0.2856	0.4258

Table 4.2.12 Rotated principal factor loadings for the genotypic correlation matrix – data B

Variable Code	Common factor coefficients		
	1	2	3
1	0.4439	0.2764	0.1379
2	0.3524	1.1433	-0.0235
4	-0.3121	-0.1728	0.9106
7	0.0302	-0.2489	0.0068
8	-0.1228	-0.2160	-0.3148
12	1.2992	-0.2284	0.0080
13	0.2515	0.1235	-0.1465
14	-0.0166	0.0084	-0.1598
15	0.0769	0.3509	0.3958
Proportionate variance accounted by each factor	0.6809	0.5319	0.3578

Factor I : number of days to maturity
height of the plant

Factor II : number of branches

Factor III : number of capsules on main stem
yield of seeds per plant

4.2.4.2 Maximum-Likelihood factor analysis

Through the Lawley's iterative scheme solutions embodying five, six, seven, eight and nine factors were successively extracted for the phenotypic correlation matrix of order 15. Some difficulty was experienced in achieving a suitable factor model for the given correlation matrix. The approximate chi-square statistics for the solutions are

Number of factors	Goodness of fit χ^2	Degree of freedom	Number of iterations
5	106.96	40	40
6	80.52	30	28
7	76.17	21	81
8	61.56	13	39
9	35.05	6	32

All these statistics lead to the rejection of the adequacy of the respective factor models at the 0.001 probability level.

The initial estimates of factor loadings and specific variances obtained from principal component method of factor analysis are given in Table 4.2.13. The five factor-model has the estimated loadings and specific variances as shown in Table 4.2.14. The unrotated and rotated loadings of the matrix are presented in Tables 4.2.15 and 4.2.16, respectively. The characters dominating in the factors are

Factor I :	number of capsules on branches number of capsules per plant yield of seeds per plant
Factor II :	height of the plant number of branches number of days to maturity height upto first capsule number of days for flowering
Factor III :	number of seeds per capsule circumference of capsule length of capsule

Table 4.2.13 Initial estimates of factor loadings and corresponding unique variances for five factors of the phenotypic correlation matrix - data B

Variable	Factor loadings					Unique variance
	1	2	3	4	5	
1	0.7329	0.2867	0.2085	0.1641	0.0186	0.6900
2	0.6546	0.2775	-0.0908	-0.4583	0.2269	0.7753
3	0.3588	0.7935	-0.0657	0.0832	0.0585	0.7730
4	0.4507	-0.6156	0.2311	0.3164	-0.0874	0.7431
5	0.8722	-0.0976	0.1342	-0.2480	0.0772	0.8557
6	0.8733	-0.3625	0.2074	-0.0420	0.0173	0.9392
7	0.1679	0.0794	0.3734	0.6988	-0.0257	0.6628
8	-0.0875	-0.0205	0.4896	0.2697	0.0775	0.3266
9	-0.3809	0.2246	0.7157	-0.1958	0.1556	0.7704
10	-0.3572	0.1720	0.7766	-0.1633	0.2186	0.8348
11	0.1679	0.7908	-0.0813	0.1333	0.2479	0.7394
12	0.2723	0.6731	-0.1087	0.2266	-0.1928	0.6276
13	0.0820	0.3057	0.1664	-0.1642	-0.8488	0.8752
14	-0.0924	0.1494	0.5143	-0.2843	-0.2643	0.4461
15	0.6812	-0.2831	0.1985	-0.0927	-0.0318	0.5932

Table 4.2.14 Maximum likelihood estimates of factor loadings and unique variances in the 39th iteration of phenotypic correlation matrix - data B

Variable	Factor loadings					Unique variance
	1	2	3	4	5	
1	0.5318	0.5946	0.0042	-0.0087	0.3291	0.2552
2	0.4921	0.1634	-0.0836	0.4839	0.1368	0.4664
3	0.0156	0.7528	-0.1289	0.4160	0.3748	0.1030
4	0.6256	0.0017	0.0014	-0.7799	0.0013	0.0003
5	0.9269	-0.0002	0.0011	0.3752	0.0006	0.0001
6	0.9967	-0.0003	-0.0006	-0.0801	-0.0005	0.0001
7	0.1464	0.1958	0.1221	-0.1617	0.1137	0.8963
8	-0.0091	-0.0044	0.2855	0.0125	0.0332	0.9172
9	-0.2193	0.0758	0.7815	0.0035	0.0402	0.3338
10	-0.1986	0.0084	0.9318	0.0018	0.1080	0.0805
11	-0.0950	0.4400	-0.0370	0.3432	0.3556	0.5517
12	0.0176	0.4831	-0.0313	0.2783	0.0559	0.6791
13	0.0071	0.5740	0.1168	0.1329	-0.7300	0.1062
14	-0.0319	0.1096	0.3301	0.0555	-0.0936	0.8662
15	0.6475	-0.0115	0.0246	0.0294	-0.0438	0.5773

Table 4.2.15 Maximum likelihood estimates of factor loadings and unique variances in the 40th iteration of phenotypic correlation matrix - data B

Variable	Factor loadings					Unique variance
	1	2	3	4	5	
1	0.5318	0.5938	0.0019	-0.0839	0.3304	0.2553
2	0.4918	0.1627	-0.0840	0.4893	0.1374	0.4664
3	0.0153	0.7514	-0.1321	0.4160	0.3769	0.1026
4	0.6262	0.0017	0.0013	-0.7795	0.0013	0.0003
5	0.9266	-0.0001	0.0011	0.3758	0.0005	0.0001
6	0.9968	-0.0003	-0.0006	-0.0795	-0.0005	0.0001
7	0.1465	0.1962	0.1212	-0.1616	0.1139	0.8863
8	-0.0091	-0.0029	0.2855	0.0124	0.0327	0.9172
9	-0.2194	0.0798	0.7808	0.0034	0.0388	0.3344
10	-0.1986	0.0130	0.9325	0.0017	0.1063	0.0795
11	-0.0952	0.4389	-0.0386	0.3432	0.3566	0.5518
12	0.0174	0.4824	-0.0835	0.2783	0.0571	0.6792
13	0.0070	0.5764	0.1124	0.1329	-0.7292	0.1057
14	-0.0319	0.1115	0.3292	0.0554	-0.0940	0.8663
15	0.6475	-0.0113	0.0247	0.0298	-0.0438	0.5773

Table 4.2.16 Rotated maximum likelihood estimates of factor loadings for the phenotypic correlation matrix - data B

Variable	Factor loadings				
	1	2	3	4	5
1	0.4885	0.4970	0.0019	-0.5081	-0.0294
2	0.6411	0.3117	-0.0840	0.1339	0.0219
3	0.1726	0.9151	-0.1321	-0.0661	-0.0915
4	0.2820	-0.4532	0.0014	-0.8449	0.0290
5	0.9999	-0.0025	0.0012	-0.0049	0.0007
6	0.8914	-0.2155	-0.0006	-0.3985	0.0134
7	0.0739	0.1021	0.1212	-0.2884	-0.0016
8	-0.0037	0.0206	0.2855	0.0069	0.0288
9	-0.2015	0.1186	0.7808	0.0341	-0.0122
10	-0.1830	0.0972	0.9325	0.0381	0.0808
11	0.0427	0.6628	-0.0386	0.0464	0.0581
12	0.1221	0.5023	-0.0835	0.0066	-0.2155
13	0.0572	0.1324	0.1124	0.0335	-0.9272
14	-0.0084	0.0672	0.3292	0.0310	-0.1407
15	0.6100	-0.1333	0.0247	-0.1779	-0.0245
Contribution of each factor	3.0255	1.2750	1.7295	2.1990	0.9480
Proportionate variance accounted by each factor	0.2017	0.0850	0.1153	0.1466	0.0632

	number of capsules on main stem
Factor IV :	number of fruiting nodes per 20 cm
Factor V :	1000 seed weight

The residual matrix formed by subtracting the correlations generated by the respective factor models from the original sample values are given in Table 4.2.17. The largest of the residual matrix had the value 0.2055.

The environment correlation matrix was subjected to ML method of factor extraction under the hypothesis that a five factor-model would suffice to explain the dependence structure. The initial estimates of loadings and specific variances were obtained by the principal component method of factor analysis. An improper maximum likelihood solution was obtained. Specific variances of variables 3, 4, 5 and 6 tend to zero in the course of iteration and become negative when allowed to continue the iteration. The ML method was applied to the partial dispersion matrix after eliminating these variables and a proper solution was obtained for one factor-model in the 24th iteration. But the goodness of fit test resulted in a Chi-square value of 87.00 for 44 degrees of freedom that leading to the rejection of the null hypothesis.

Table 4.1.17 Residual matrix after removal of five factors, phenotypic correlation matrix - data B

0														
-0.0118	0													
0.0071	0.0067	0												
0.0005	-0.0007	0.00008	0											
0.0003	-0.0005	0.00006	-0.0000	0										
-0.0003	0.0004	-0.0000	0.0000	0.0000	.0									
-0.0616	-0.1111	0.0084	0.0012	0.0008	-0.0006	0								
0.0232	-0.1267	0.0096	0.0011	0.0007	-0.0006	0.1526	0							
-0.0356	-0.0084	0.0115	-0.0006	-0.0004	0.0003	-0.0142	-0.0596	0						
0.0062	0.0073	-0.0029	0.0001	0.0000	-0.0000	-0.0032	0.0065	0.0035	0					
-0.0513	0.0453	-0.0032	-0.0015	-0.0010	0.0007	0.0913	-0.0945	0.0239	0.0013	0				
0.0265	-0.0583	-0.0324	-0.0006	-0.0004	0.0003	0.0681	-0.0368	-0.0239	0.0051	0.2055	0			
-0.0021	0.0024	0.0009	-0.0000	0.0000	0.0000	-0.0015	-0.0013	0.0008	0.0000	-0.0071	0.0070	0		
0.0342	0.0077	-0.0070	0.0002	0.0001	-0.0001	0.0376	0.0142	0.0011	-0.0032	0.0141	-0.1027	0.0039	0	
0.0554	-0.0139	-0.0122	0.0000	0.0000	0.0000	0.0051	0.0455	-0.0545	0.0086	-0.0378	0.0018	0.0013	-0.0249	0

Therefore ML analysis was repeated with six factors. After eleven iterations the 3rd, 4th, 5th and 6th specific variances become negative and the analysis was done for the partial dispersion matrix after eliminating these variables. After six iterations a proper solution was obtained for the two factors. The value of the test criterion was found to well below expectation and the hypothesis was accepted. The initial estimates of factor loadings and unique variances for two factors (for environment correlation matrix of order 11 after eliminating variables 3, 4, 5 and 6) are given in Table 4.2.18. The maximum likelihood estimates of factor loadings and corresponding unique variances for the six factor-model (combined solution) are given in Table 4.2.19. The rotated loadings are presented in Table 4.2.20. The variables dominating the factors are

Factor I :	number of capsules on main stem
	height upto first capsule
	number of capsules on branches
	number of capsules per plant
	yield of seeds per plant
Factor II :	height of the plant

Table 4.2.18 Initial estimates of factor loadings and unique variances for the environment correlation matrix of order 11, two factor case - data B

Variable Code	Factor loadings 1	Factor loadings 2	Unique variance
1	0.0107	-0.0192	0.9995
2	-0.1771	-0.1619	0.9424
7	-0.3107	0.1450	0.8824
8	0.3234	0.6797	0.4334
9	0.7127	-0.0356	0.4908
10	0.6430	0.5112	0.3252
11	-0.4588	0.3851	0.6412
12	-0.3839	0.4731	0.6288
13	-0.3501	0.4435	0.6807
14	0.0472	-0.0428	0.9959
15	0.0742	0.1223	0.9757

Table 4.2.19 Maximum likelihood estimates of factor loadings and unique variances for the environment correlation matrix, six factor - model (combined solution) - data B

Variable	Factor loadings						Unique variance
	1	2	3	4	5		
1	0.3261	0.9220	0.1551	0	0	0.0004	0.0195
2	0.7821	-0.2650	0.5600	0	0	0.0049	0.0045
3	0.8922	0.0123	-0.4431	0	0	0.0094	0.0074
4	0.9889	-0.1076	-0.0891	0	0	-0.0123	0.0024
5	0.7225	0.2546	0.2268	-0.04	-0.0248	0.0470	0.3573
6	0.4404	0.2151	-0.2657	-0.1247	-0.0170	-0.0503	0.6708
7	0.3478	0.1232	0.2743	-0.1124	0.1801	-0.0057	0.7435
8	0.2876	-0.0520	0.1163	0.3939	0.2761	0.0790	0.6634
9	0.2563	-0.0579	-0.0586	0.4619	-0.2930	0.0849	0.6211
10	0.0953	-0.1288	-0.0493	0.7299	0.1302	0.0319	0.4212
11	0.0157	0.1126	-0.1706	-0.1336	0.3762	-0.1804	0.7661
12	0.1926	0.1467	-0.2276	-0.0876	0.3914	0.0437	0.7268
13	0.0223	0.0993	0.1507	-0.1364	0.3080	0.0296	0.8526
14	0.0008	0.1555	0.0256	-0.0279	-0.0417	-0.0024	0.9773
15	0.6372	-0.0728	-0.1199	0.0656	0.0492	-0.1133	0.5547

Table 4.2.20 Rotated maximum likelihood solution for the environment correlation matrix - data B

Variable	Factor loadings					
	1	2	3	4	5	6
1	0.3261	0.9241	-0.0401	0.0626	0.1829	0.0004
2	0.7821	-0.1417	0.6035	-0.0096	-0.0281	0.0049
3	0.8922	-0.0789	-0.4437	-0.0054	-0.0156	0.0094
4	0.9889	-0.1221	-0.0649	-0.0083	-0.0242	-0.0123
5	0.7225	0.2948	0.1692	-0.0261	0.0473	0.0470
6	0.4404	0.1557	-0.3045	-0.1131	0.0548	-0.0503
7	0.3478	0.1366	0.2429	-0.0375	0.2373	-0.0057
8	0.2876	-0.0828	0.1245	0.4584	0.1229	0.0790
9	0.2563	-0.0073	-0.0454	0.3396	-0.4342	0.0849
10	0.0953	-0.1599	-0.0216	0.7228	-0.1422	0.0319
11	0.0157	-0.0037	-0.1902	-0.0022	0.4062	-0.1804
12	0.1926	0.0143	-0.2530	0.0476	0.4095	0.0437
13	0.0223	0.0625	0.1269	-0.0229	0.3543	0.0296
14	0.0008	0.1626	-0.0071	-0.0292	0.0009	-0.0024
15	0.6372	-0.1041	-0.1023	0.0713	0.0057	-0.1133
Contribution of each factor	3.9300	1.0995	0.8925	0.8760	0.7785	0.6765
Proportionate variance accounted by each factor	0.2620	0.0733	0.0595	0.0584	0.0519	0.0451

Factor III :	number of branches
	number of fruiting nodes per 20 cm
	number of seeds per capsule
Factor IV :	length of capsule
	circumference of capsule
Factor V :	number of days to maturity
	number of days to flowering
Factor VI :	1000 seed weight

The residual correlation matrix of order 11 after elimination of two factors is given in Table 4.2.21. The elements of the matrix are considerably small.

The positive definite genotypic correlation matrix of order nine was analysed using the Lawley's iterative scheme method and maximum likelihood solutions were successively extracted for three, four and five factor-models. Details of the goodness-of-fit tests for the three solutions are summarised as

Number of Factors	Chi-square	Degrees of freedom	Number of iterations
3	79.83	12	142
4	68.19	6	25
5	26.09	1	40

Table 4 2.21 Residual matrix after removal of six factors, environment correlation matrix - data B

0										
-0.0061	0									
-0.0243	0.0228	0								
0.1178	-0.0449	0.0082	0							
0.0424	0.0029	-0.0404	0.0092	0						
-0.0773	-0.0099	0.0027	0.0338	0.0260	0					
0.0434	0.1052	0.1378	0.0094	-0.0351	0.0166	0				
-0.0309	-0.0122	-0.0442	0.0025	-0.0517	0.0241	0.0141	0			
0.0576	-0.0938	-0.0736	0.0707	0.0225	-0.0486	-0.0535	0.0915	0		
0.0388	-0.0587	-0.1372	0.0588	0.0035	-0.0384	-0.0392	0.0593	-0.0073	0	
0.0645	-0.0077	-0.0446	0.0999	-0.0755	0.0114	0.0245	-0.1250	0.02119	0.0228	0

All these statistics lead to the rejection of the adequacy of the respective factor models at the 0.001 probability level. The initial estimates of loadings and specific variances are presented in Table 4.2.22. The three-factor model had the estimated loadings and specific variances obtained respectively in the 141st and 142nd iterations are as shown in Tables 4.2.23 & 4.2.24. The rotated factors are given in Table 4.2.25. The variables influencing the factors are

Factor I : number of branches

Factor II : number of capsules on main stem
yield of seeds per plant

Factor III : number of days to maturity
height of the plant

The residual correlations after the extraction of five-factors are presented in Table 4.2.26.

Table 4.2.22 Initial estimates of factor loadings and corresponding unique variances for three factors of the genotypic correlation matrix - data B

Variable Code	Factor loadings			Unique variance
	1	2	3	
1	0.7186	-0.0269	0.2375	0.4264
2	0.7977	-0.1936	-0.1870	0.2912
4	-0.3821	-0.6842	0.4227	0.2072
7	-0.1983	0.2902	0.7726	0.2796
8	-0.3568	0.5691	0.0824	0.5420
12	0.6582	0.3442	0.1012	0.4380
13	0.4691	0.3404	0.1774	0.6326
14	0.0412	0.3948	0.3467	0.7222
15	0.3863	-0.5943	0.4103	0.3292

Table 4.2.23 Maximum likelihood estimates of factor loadings and unique variances in the 141^{th} iteration of genotypic correlation matrix of order 9 - data B

Variable Code	Factor loadings			Unique variance
	1	2	3	
1	0.4905	-0.1242	0.5165	0.4773
2	0.9372	-0.3280	-0.0078	0.0140
4	-0.5997	-0.7234	0.3040	0.0246
7	-0.2485	0.1766	0.1364	0.8884
8	-0.1512	0.2851	-0.2550	0.8308
12	0.3892	0.4895	0.7540	0.0405
13	0.2778	0.1893	0.1378	0.8680
14	0.0150	0.0541	-0.1401	0.9772
15	0.2664	-0.5337	0.2023	0.6032

Table 4.2.24 Maximum likelihood estimates of factor loadings and unique variances in the 142nd iteration of genotypic correlation matrix of order 9 - data B

Variable Code	Factor loadings			Unique variance
	1	2	3	
1	0.4906	-0.1236	0.5165	0.4773
2	0.9373	-0.3277	-0.0076	0.0139
4	-0.5994	-0.7234	0.3047	0.0245
7	-0.2485	0.1766	0.1363	0.8885
8	-0.1513	0.2848	-0.2552	0.8309
12	0.3891	0.4903	0.7537	0.0402
13	0.2778	0.1895	0.1376	0.8680
14	0.0149	0.0540	-0.1402	0.9772
15	0.2666	-0.5334	0.2028	0.6033

Table 4.2.25 Rotated maximum likelihood estimates of factor loadings of the genotypic correlation matrix - data B

Variable Code	Factor loadings		
	1	2	3
1	0.3600	-0.1236	0.5632
2	0.8712	-0.4278	0.2459
4	-0.6700	-0.7234	0.0566
7	-0.2816	0.1766	0.0327
8	-0.0441	0.2934	-0.2848
12	0.1764	0.5903	0.6447
13	0.2055	0.1894	0.2321
14	0.0666	0.0540	-0.1242
15	0.1706	-0.5334	0.2882
Contribution of each factor	1.5256	1.5106	1.0310
Proportionate variance accounted by each factor	0.1695	0.1678	0.1146

Table 4.2.26 Residual matrix after removal of three factors

0								
0.0030	0							
0.0038	-0.0001	0						
0.0896	-0.0026	-0.0026	0					
0.0799	-0.0004	-0.0015	0.1632	0				
0.0028	-0.0001	-0.0002	-0.0030	-0.0045	0			
-0.1159	-0.0003	-0.0005	0.0112	-0.0222	0.0055	0		
0.1527	-0.0013	-0.0002	0.1159	0.0700	-0.0101	0.2437	0	
-0.1356	0.0026	0.0028	0.1617	0.0128	0.0035	0.1491	-0.0292	0

DISCUSSION

DISCUSSION

Multivariate statistical methods are used increasingly in plant breeding research to investigate the responses of plants of different genetic origin. The relevance of these methods is in seeing the plants as a whole rather than individual ones as the plant breeder wishes. The plants exhibit dependence structure since the performance of a plant depends upon various morphological and quality traits. In addition, these methods have had their chief success in plant breeding where the problems clearly lie in the integration of numerous related traits ie, in defining the dependence structure of the variables considered. This dependence structure is resolved into their putative underlying causes using factor analytic methods.

One of the final outcome of factor analysis is the factor loading matrix. The factor loadings are the component correlations between the response variable and the factor. Once the factor loadings have been extracted then the problem is to make the best interpretation of the common factors. Varimax rotation helps to make the best interpretation of these common factors. Each factor in the factor loading matrix is dominated by the variables with large loadings in absolute value and these variables are highly correlated with the respective factor.

Moreover, seed yield in sesame is a complex character contributed by a number of intercorrelated traits. Information on the extent of genetic diversity will be of much use to a plant breeder in further breeding programmes. The strategy of plant breeding relies to a great extent upon a proper programme of hybridisation, which in turn seeks a choice of potential parents. Factor analytic techniques provide supplimentary information on the diversity with a lesser number of causative factors. It is a well known fact that the heterosis obtained in a cross between genetically diverse parents will be large.

The causative sources of common variation within the species are analysed, using the principal factor and maximum-likelihood methods of factor analysis. The analysis of variance for each character show significant differences for all the characters except number of fruiting nodes per 20 cm studied under the two environments - upland and rice fallows, indicating the need for further genetic analysis. Wilk's criterion also reveal highly significant differences among the varieties for the aggregate of all characters, indicating that the varieties differ significantly from each other when all the characters are considered simultaneously. In the light of information provided by the eigen structure of the correlation matrices it is decided to represent the data by a five-factor model. The choice of the number of common factors is motivated by the fact of taking as many factors as there are whose eigen values are greater than or equal to one (Guttman, 1956; Afifi and Azen, 1972). Sometimes the experimenter may satisfied with a smaller number of factors. It is known that the number of factors extracted, will depend upon the material taken for investigation (Cattell, 1965 a). The number of common factors taken up for interpretation in factor analysis generally depends upon their aggregative explanatory power and theoretical approach of a factor analyst (Shetty, 1969).

The characters which are highly correlated with each factor are identified by varimax rotation. Each variable may reasonably be assigned to that factor with which it shows the closest linear relationship ie, that factor in which it has the highest loading. Accordingly, the variables which are closely related are clustered together in that factor. Shetty (1969) used the same method of clustering in each factor. The factors are named from the nature and magnitude of the variables which they represent. A factor is named from the common attribute of these variables regardless of the specific content (Harman, 1967).

In addition to the factor weights, the proportionate variance accounted by each factor in terms of total original communality is important in the case of principal-factor solution. The general characteristic of the principal-factor solution is that the contribution of the factors to the total original communality of the variables decreases with each succeeding factor. The contributions of the factors and total communality is of importance in maximum likelihood solutions, while maximising variance of each factor is important in principal-factor solution (Harman, 1967).

5.1 Phenotypic dependence structure

The five factors identified for phenotypic correlation matrix in uplands by PF method are as follows. The first factor is highly correlated with number of capsules per plant, number of capsules on main stem, number of capsules on branches and yield of seeds per plant. These variables are related to the reproduction of the crop and hence named as reproductive factor. The number of branches forms the second factor which is the vegetative factor. Third factor is the height factor consists of height of the plant and height upto first capsule. The seed characters like circumference of capsule, length of capsule and number of seeds per capsule forms the fourth factor which are related to the characteristic of seeds and hence called as the seed factor. Number of days to maturity and number of days to flowering constitute the fifth factor and this factor is identified as growth factor as the variables are related to the growth of the plant.

ML method leads to an eight factor model from the point of view of goodness of fit. The first five factors are same as that of above but with a change in the order of factors. The sixth and seventh factors are characterised by 1000 seed weight and seed oil content which are called respectively as weight factor and quality factor as they are related to the quantity and quality of seed. The eighth factor is the growth factor which is correlated with number of days to flowering. The number of days to flowering and number of days to maturity are characters related to growth of the crop but in this case they are identified as two different factors.

The comparison of the results shows that the first and third factor extracted by the two methods are same viz., reproductive factor and height factor. In the ML method seed character has more importance than vegetative character as in PF method. The additional factors identified by ML method are weight factor and quality factor.

Reproductive factor is identified as the first factor in rice fallows by PF method. The second factor is associated with height upto first capsule, number of days to maturity, number of days to flowering, height of the plant and number of branches. All these variables are related to the growth of the crop and hence come under the growth factor. The third, fourth and fifth factors are identified respectively as seed factor, weight factor and density factor.

The ML method does not give a good-fit to the factor model even after fixing the upper limit to the number of factors. So the five factor model ie, the model with the minimum number of factors is considered for the comparison purpose. This model identified the same factors as given by PF method but with a change in the magnitude of variation explained by the last two factors. The fourth factor in the PF method is the fifth factor in ML method and vice versa.

Principal-factor solution with SMC's as estimates of the communalities produced five factors. They contributed 92.44 percent and 109.62 percent to total original communality respectively in uplands and rice fallows. The total communality produced by the maximum-likelihood solution represents 75.96 percent and 61.18 percent of the total variance of the variables respectively in uplands and rice fallows. The first factor amounts to twice the variability explained by the second factor except using PF solution in rice fallows. The variability explained by each factor decreases with the additive inclusion of factors in succession, which is also a property of the methods adopted.

Yield of seeds per plant, number of capsules per plant, number of capsules on branches are highly correlated with the first factor in both the situations under the two different methods. The number of capsules per plant is found to have maximum loading with first factor but the estimate of communality was larger in magnitude than the respective loading. Seed yield has also got high loading with first factor but with a low communality. Sawanth et al. (1982) has reported that the characters in a given factor with which high factor loading and low communality is amenable to change very easily due to selection as compared to one with high factor loading and high communality. So seed yield may be considered as the important character in factor one.

The character 'height upto first capsule' is more amenable to change in the height factor than height of the plant. The three seed characteristic variables viz., the circumference of capsule, length of capsule and number seeds per capsule are of equal importance while using seed factor. The number of days to maturity has more importance in growth factor in uplands and number of days to flowering and number of days to maturity have equal importance in rice fallows. Number of fruiting nodes per 20 cm has importance while selecting density factor in rice fallows.

5.2 Environment dependence structure

The environment correlation matrix in uplands is explained by five factors using PF method. The first, second and third factors are respectively as reproductive, height and seed factors. The fourth factor is dominated by number of days to maturity which is a growth factor. The number of fruiting nodes per 20 cm and number of branches constitute the fifth factor which is termed as density factor as the characters are related to the density of the crop.

Initiating from a five factor model an eight factor solution is found to fit the matrix by ML method. The first three factors are similar to that obtained by PF method. Growth factor is also identified but its order is changed. The density factor obtained in the PF method is split into two separate factors namely vegetative factor consisting of number of branches and density factor dominated with number of fruiting nodes per 20 cm only. The reproductive factor in PF method is identified as two factors in ML method. The number of capsules on branches forms a separate factor in the eight factor model. The only additional factor identified by ML method is the quality factor which is concerned with the oil content of the seed.

The five-factor model fitted to rice fallows by PF method is in agreement with that of uplands. The last two factors are found to interchange their order of occurrence. ML method helps to set a satisfactory solution to six-factor model for the environment correlation matrix in rice fallows. The first five factors are same as that obtained by PF analysis, the magnitude of variation being different. The additional factor identified is the weight factor which is dominated by 1000 seed weight. The growth factor identified in rice fallows includes number of days to flowering and number of days to maturity while in uplands it consists of only number of days to maturity.

The PF analysis is found to give a better comparison of the factors identified for the environment correlation matrix in two situations. It reveals that five factors viz., reproductive, height, seed, growth and density factors are the underlying causes of diversity in sesame plants. But ML method has got its own properties. One useful advantage of ML method used in the present study as compared to PF analysis is that estimates are scale invariant. Secondly an adequate number of factors for better explanation of original data is obtained. ML analysis resulted in the identification of a quality factor in uplands and weight factor in rice fallows in addition to the five factors identified by PF method for the adequate representation of dependence structure of original data.

The five-factor model obtained by the PF analysis contributed 105.44 percent and 102.08 percent to total original communality respectively in uplands and rice fallows. The total communality produced by the ML solution represents 68.48 percent and 55 percent of total variance of the variables respectively in uplands and rice fallows.

The characters which are more amenable to change due to selection in uplands are yield of seeds per plant in reproductive factor, height upto first capsule in height factor, circumference of capsule in seed factor, number of days to maturity in growth factor and number of fruiting nodes per 20 cm in density factor. But in rice fallows, the characters identified are number of capsules on branches and yield of seeds per plant in reproductive factor, height upto first capsule in height factor, circumference of capsule and number of seeds per capsule in seed factor, number of fruiting nodes per 20 cm in density factor, number of days to maturity in growth factor.

High communality obtained in almost all cases indicates the high reliability of the results that are obtained (Shetty, 1969). Similar results are obtained in the present study also. Environment correlation matrices have given same factor pattern in different situations under study, while phenotypic and genotypic correlation matrices fails to give same factor pattern in two environments. So, the environment correlation matrix is found to be the appropriate estimate of population correlation matrix and can be used for factor analytic studies. Similar results are reported by Murty and Arunachalam (1967) and Muralidharan (1986).

5.3 Genotypic dependence structure

In the present study, during both rabi and summer, genotypic correlations are found slightly higher than the respective phenotypic correlation coefficients. This indicates the masking effect of the environment to the total expression of the genotypes. This may be due to the high experimental error detected in the conduct of the experiment. These results are in accordance with the reports of Dabral (1967); Sanjeeviah and Joshi (1974); Thangavelu and Rajasekharan (1983) and Sverup John (1985).

All correlation matrices with unities in the principal diagonal are Gramian matrices - matrix with Gramian properties (Harman, 1967). Here the genotypic correlation matrix of order 15 is found to be in indefinite form under both the situations. This may be due to the fact that they are not estimated by product-moment methods. Kendall (1983) gave a warning against the attempts at component or factor analysis of matrices which are not obtained by product-moment methods. The correlation matrices estimated by other methods may not necessarily positive definite, and in certain cases some of the latent roots may turn out to be negative. Trial and error method is used to make the genotypic correlation matrix in positive definite form. The variables with high genotypic correlation coefficients are eliminated one by one and the resulting matrix is tested for positive definiteness. The process is repeated till the matrix reached in positive definite form. Genotypic correlation matrices of order 10 and 9 are thus obtained respectively in uplands and rice fallows. Similar results were reported for the genotypic correlation matrices under different environments in a study on genetic divergence of groundnut varieties (Muralidharan, 1986).

Three factors are identified with PF analysis. The first factor consist of number of fruiting nodes per 20 cm, length of capsule and circumference of capsule. This is a combination of seed and density characters. The second factor is associated with height and vegetative character and third factor is concerned with growth and quality.

First factor is same in both ML and PF methods. But the second factor in ML is blended with height and yield. The number of branches and oil content form the third factor.

In the rice fallows, first factor consisted of number of days to maturity and height of the plant. These characters are associated with growth factor. The number of branches forms the second factor and third factor is blended with number of capsules on main stem and yield of seeds per plant. This is a reproductive factor. ML analysis also identified the same factors with a change in the contribution of variance.

In both situations the ML analysis failed to give a satisfactory fit to the reduced genotypic correlation matrix. Another fact to be noted is that the same pattern of factors is obtained even after varimax rotation. Since the genotypic correlation matrices in two cases are of different order and the pattern of factor solution is also different, a comparative study is not worth. Genotypic correlation matrices are not suitable for this type of analysis as they do not possess the properties of Wishart distribution. In the present study the 100 varieties were selected from 252 varieties based on the general performance of plants. This may be one reason for getting genotypic correlations highly skewed. Murty and Arunachalam (1967) has reported that directional selection by man resulted in highly skewed genetic correlations. He also pointed out that genotypic variance-covariance matrix need not necessarily be an estimate of the parameter of a multivariate normal distribution.

SUMMARY

SUMMARY

Multivariate statistical techniques are widely used in plant breeding research to estimate the degree of divergence in morphological and quality traits as they are intercorrelated to varying degrees. Factor analysis is considered as the queen of analytic methods due to its power and elegance in studies of this type. Principal factor method and maximum likelihood method are two ways to extracting the factors of divergence, of which maximum likelihood method is considered as the best one as it satisfies certain properties of a best estimator. Also it allows for the determination of an adequate number of stable factors from the point of view of goodness of fit of the factor-model.

The available data on various morphological characters and oil content in sesame with respect to hundred varieties grown in upland and rice fallow in 1981-'82 were utilized for the study. The MANOVA revealed significant differences among the varieties for aggregate effect of all the above characters indicating considerable variability among the experimental material. The various factor-models were tried for the phenotypic, environment and genotypic correlation matrices as factor analysis aims

to explain the intercorrelations among the numerous variables in terms of simpler relations.

Principal factor analysis allows for the determination of a m-factor pattern where m refers to the number of principal components whose eigen values are greater than or equal to one (Harman, 1967). As such a five-factor model was fitted to the phenotypic and environment correlation matrices under both situations. The five factors identified for phenotypic correlation matrix in uplands were reproductive, vegetative, height, seed and growth factors. In rice fallows they are reproductive, growth, seed, weight and density factors. The maximum likelihood method resulted in the fitting of an eight-factor model in uplands and fails to give an adequate factor-model in the rice fallows. The additional factors identified in uplands were weight and quality factors. In both situations the first factor identified was reproductive factor. The factor pattern identified in the two environments differ slightly.

The analysis of environment correlation matrix by principal factor method identified five factors viz., reproductive, height, seed, density and growth factors for the parsimonious summarisation of the data under the two environments. The maximum likelihood analysis revealed that an additional quality factor was working in uplands and a weight factor in rice fallows. It shows an adequate fit of an eight-factor model in uplands and a six-factor model in rice fallows.

The characters which were more amenable to change due to selection in uplands were yield of seeds per plant in reproductive factor, height upto first capsule in height factor, circumference of capsule in seed factor, number of days to maturity in growth factor and number of fruiting nodes per 20 cm in density factor. In rice fallows the same characters were identified in height, growth and density factors. The characters number of capsules on branches and yield of seeds per plant were identified in reproductive factor, circumference of capsule and number of seeds per capsule in seed factor.

The genotypic correlation matrices under the two environments were found not suitable for factor analytic studies as it lacks properties of this type of analysis. The environment correlation matrix was found to be appropriate for factor analytic studies as it gives stable factor pattern under two environments.

REFERENCES

REFERENCES

Abraham, T.P. and Hoobakht, A. (1974). An application of factor analysis for interpretation of soil analysis data. J. Indian Soc. Agrl. Stat. 26(1): 103-112.

Afifi, A.A. and Azen, S.P. (1972). Statistical Analysis a Computer Oriented Approach. Academic Press, Inc. New York. pp. 361.

Anderson, T.W. (1958). An Introduction to Multivariate Statistical Analysis. John Wiley and Sons, Inc. New York. pp. 374.

*Bartlett, M.S. (1934). The vector representation of a sample. Proc. Camb. Phil. Soc. 30: 327-340.

Bartlett, M.S. (1947). Multivariate Analysis. J. Roy. Stat. Soc. (B) 9: 176-197.

Bartual, B.; Carbonell, E.A. and Green, D.E. (1985). Multivariate analysis of a collection of soybean cultivars for southwestern Spain. Euphytica 34(1): 113-123.

*Burt, C. (1941). The factors of the mind : an introduction to factor analysis in psychology. Macmillan. New York.

Burt, C. (1952). Tests of significance in factor analysis. Brit. J. Psychol. Stat. Sec. 5: 109-133.

Cattell, R.B. (1965 a). Factor analysis : An introduction to essentials 1. The purpose and underlying models. Biometrics. 21(1): 190-215.

Cattell, R.B. (1965 b). Factor analysis : An introduction to essentials II. The role of factor analysis in research. Biometrics. 21(2): 405-435.

Cattell, R.B. and Khanna, D.K. (1977). Principles and procedures for unique rotation in factor analysis. Mathematical Methods for Digital Computers. 3. Ed. Enslein, K.; Ralston, A. and Wilf, H.S. John Wiley and Sons. Inc. New York.

Clarke, M.R.B. (1970). A rapidly convergent method for maximum-likelihood factor analysis. British J. Stat. Psych. 23: 43-52.

Cochran, W.G. and Cox, G.M. (1957). Experimental designs. 2nd ed. John Wiley and Sons, Inc. London. pp. 611.

Dabral, K.C. (1967). Variability and correlation studies in sesame. JNKVV Res. J. 1: 135-139.

Denis, J.S. and Adams, M.W. (1978). A factor analysis of plant variables related to yield in dry beans. Crop. Sci. 18(1): 74-78.

Federer, W.T. (1955). Experimental design. Macmillan Company. New York. pp. 544.

*Fisher, R.A. (1922). On the mathematical foundations of theoretical statistics. Phil. Trans. Roy. Soc., 222: 309-368.

*Fisher, R.A. (1925). Theory of statistical estimation. Proc. Camb. Phil. Soc. 22: 700-725.

Frutcher, B. (1954). Introduction to Factor Analysis. D. Van Nostrand Company Inc. New York. pp. 280.

*Garnett, J.C.M. (1919). On certain independent factors in mental measurement. Proc. Roy Soc. London. 46: 91-111.

Gaur, P.C.; Gupta, P.K. and Kishore, H. (1978). Studies on genetic divergence in potato. Euphytica 27: 361-368.

*Gill, R.D. (1977). Consistency of maximum likelihood solutions of factor analysis model, when the observations are not multivariate normal. Recent Developments in Statistics. Ed. Barra, J.R. et al., North-Holland.

Guttman, L. (1956). "Best possible" systematic estimates of communalities. Psychometrika 21(3): 273-285.

Harman, H.H. (1960). Modern Factor Analysis. University of Chicago Press, Chicago. pp. 471.

Harman, H.H. (1967). Modern Factor Analysis. 2nd ed. University of Chicago Press, Chicago. pp. 473.

Hemmerle, W.J. (1965). Obtaining maximum-likelihood estimates of factor loadings and communalities using an easily implemented iterative Computer Procedure. Psychometrika. 30(3): 291-302.

*Holzinger, K.J. and Harman, H.H. (1941). Factor Analysis. University of Chicago Press, Chicago.

Hotelling, H. (1931). The generalization of Student's ratio. Ann. Math. Stat. 2: 360.

Hotelling, H. (1933 a). Analysis of a complex of statistical variables into principal components. J. Educ. Psychol. 24: 417-441.

Hotelling, H. (1933 b). Analysis of a complex of statistical variables into principal components. J. Educ. Psychol. 24: 498-520.

*Hotelling, H. (1935). The most predictable criterion. J. Educ. Psychol. 26: 139-142.

Hotelling, H. (1936). The relation between two sets of varieties. Biometrics 28: 321-377.

Hussaini, S.H.; Goodman, M.M. and Timothy, D.H. (1977). Multivariate analysis and the geographical distribution of the world collection of Finger millet. Crop Sci. 17(2): 257-263.

*Jennrich, R.I. and Thayer, D.T. (1973). A note on Lawley's formula for standard errors in maximum likelihood factor analysis. Psychometrika 38: 571.

Jeswani, L.M.; Murty, B.R. and Mehra, R.B. (1970). Divergence in relation to geographical origin in a world collection of Linseed. Indian J. Genet. 30(1): 11-25.

Joreskog, K.G. (1967). Some contributions to maximum likelihood factor analysis. Psychometrika. 32(4): 443-482.

Joreskog, K.G. (1969). A general approach to confirmatory maximum likelihood factor analysis. Psychometrika 34(2): 183-202.

Joreskog, K.G. (1971). Simultaneous factor analysis in several populations. Psychometrika 36: 409.

Joreskog, K.G. (1977). Factor analysis by least-squares and maximum-likelihood methods. Mathematical Methods for Digital Computers. 3. Ed. Enslein, K.; Ralston, A. and Wilf, H.S. John Wiley and Sons, Inc. New York.

Joreskog, K.G. and Goldberger, A.S. (1972). Factor analysis by generalized least squares. Psychometrika 37:

*Kaiser, H.F. (1956). The varimax method of factor analysis. Unpublished Ph.D. Thesis, University of California.

Kaiser, H.F. (1958). The varimax criterion for analytical rotation in factor analysis. Psychometrika 23: 187-200.

*Kaiser, H.F. (1960). The application of electronic computers to factor analysis. Edc. Psychl. Measurement. 20: 141-151.

Kaiser, H.F. and Caffrey, J. (1965). Alpha factor analysis. Psychometrika 30: 1-14.

Kamboj, R.K. and Mani, S.C. (1983). Genetic diversity in triticale. Indian J. Genet. 43(2): 173-179.

Kandaswamy, M.K. (1985). Genetic variation and Genotypic environment interaction in sesamum. Madras Agric. J. 72(3): 156-161.

*Kelley, T.L. (1935). Essential traits of mental life. Harward Studies in Education. 26 Harvard University Press, Cambridge.

Kendall, M.G. (1950). Factor analysis as a statistical technique. J. Roy. Stat. Soc. (B) 12: 60-73.

Kendall, M.; Stuart, A. and Ord, J.K. (1983). The Advanced Theory of Statistics. 3 (4th ed.) Charles Griffin and Company. London.

Krishnadoss, D. and Kadambavana Sundaram, M. (1986). Correlation between yield and yield components in sesame. J. Oilseeds. Res. 3(2): 205-209.

Kruskal, J.B. and Shepard, R.N. (1974). A non metric variety of linear factor analysis. Psychometrika 39: 123.

Kukadia, M.U. and Singhania, D.L. (1984). Factor analysis in sorghum for forage yield. Indian J. Agric. Sci. 54(11): 1001-1003.

*Lawley, D.N. (1940). The estimation of factor loadings by the method of maximum likelihood. Proc. Roy. Soc. Edin. 60: 64-82.

*Lawley, D.N. (1942). Further investigations in factor estimation. Proc. Roy. Soc. Edin. 61: 176-185.

Lawley, D.N. (1943). The application of the maximum likelihood method for factor analysis. Brit. J. Psych. 33: 172-175.

Lawley, D.N. and Maxwell, A.E. (1963). Factor Analysis as a Statistical Method. Butterworth, London. pp. 177.

Lawley, D.N. and Maxwell, A.E. (1971). Factor Analysis as a Statistical Method. (2nd ed.) Butterworth, London. pp. 148.

Martin, N.G. and Eaves, L.J. (1977). The genetical analysis of covariance structure. Heridity 38(1): 79-95.

*Maxwell, A.E. (1961). Recent trends in factor analysis. J. Roy. Stat. Soc. (A) 124: 49-59.

McDonald, R.P. (1970). The theoretical foundations of principal factor analysis, Canonical factor analysis and alpha factor analysis. Brit. J. Math. Stat. Psychol. 23(1): 1-21.

Morrison, D.F. (1978). Multivariate Statistical Methods. 2nd ed. McGraw-Hill, Inc. pp. 415.

Mulaik, S.A. (1972). The foundations of factor analysis. McGraw-Hill, Inc. New York. pp. 453.

Muralidharan, K. (1986). Assessment of genetic divergence by factor analysis in groundnut. Unpublished M.Sc.(Ag.) Stat. Thesis, K.A.U.

Murty, B.R. and Anand, I.J. (1966). Combining ability and genetic diversity in some varieties of linum usitatissimum. Indian J. Genet. 26(1): 21-36.

Murty, B.R. and Arunachalam, V. (1967). Factor analysis of diversity in the Genus Sorghum. Indian J. Genet. 27(1): 123-135.

Murugesan, M.; Dhamu, K.P. and Arokia Raj, A. (1979). Genetic variability is some quantitative characters of sesamum. Madras Agric. J. 66(6): 366-369.

*Neyman, J. and Pearson, E.S. (1928). On the use and interpretation of certain test criteria for purpose of statistical inference. Biometrics 20(A): 175.

*Neyman, J. and Pearson, E.S. (1931). On the problem of K samples. Bull. Int. Acad. Cracovie (A): 460.

Paramasivam, K. and Prasad, M.N. (1980). Character association analysis in sesamum crosses. Madras Agric. J. 67(11): 701-705.

*Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. Philosophical Magazine (series 6) 2: 559-572.

*Pearson, E.S. and Neyman, J. (1930). On the problem of two samples. Bull. Int. Acad. Cracovie A: 73.

Ram, J. and Panwar, D.V.S. (1970). Intranspecific divergence in rice. Indian J. Genet. 30(1): 1-10.

Rao, C.R. (1952). Advanced Statistical Methods in Biometric Research. John Wiley and Sons, Inc. New York.

Rao, C.R. (1955). Estimation and tests of significance in factor analysis. Psychometrika 20(2): 93-111.

*Reddy, M.B. (1981). Studies on genetic variability, heterosis, combining ability, gene action and character associations in sesamum crosses. Unpublished Ph.D. Thesis, A.P.A.U., Hyderabad.

Reddy, M.B.; Reddi, M.V. and Rana, B.S. (1984). Combining ability studies in sesame. Indian J. Genet. 44(2): 314-318.

Roff, M. (1936). Some properties of the communality in multiple factor theory. Psychometrika 1: 1-6.

Sanjeeviah, B.S. and Joshi, M.S. (1974). Correlation and genetic variability in sesamum. Curr. Res. 11: 144-145.

Sawant, A.R.; Asawa, B.M. and Rawat, G.S. (1982). Factor analysis of ninety diversified strains of triticale. Indian J. Agric. Sci. 52(4): 209-211.

Schilderinck, J.H.F. (1978). Regression and Factor analysis in econometrics. 2nd ed. 1: pp. 236.

Shetty, N.S. (1969). A factor analysis of use of fertilizers by farmers. Indian J. Agril. Eco. 24: 50-61.

Singh, K.P.; Kakar, S.N.; Singh, V.P. and Chaudary, B.D. (1982). Genetic divergence in chickpea. Indian J. Agric. Sci. 52(8): 492-495.

Somayajulu, P.L.N.; Joshi, A.B. and Murty, B.R. (1970). Genetic divergence in wheat. Indian J. Genet. 30: 47-58.

*Spearman, C. (1904). General intelligence objectively determined and measured. Amer. J. Psychol. 15: 201-293.

Sundaram, A.; Ramakrishnan, A.; Marappan, P.V. and Balakrishnan, G. (1980). Factor analysis in Cowpea. Indian J. Agric. Sci. 50(3): 216-218.

Sverup John (1985). Genetic variability, path analysis and stability parameters in sesame. Ph.D. (Ag.) Thesis, K.A.U.

Swain, A.J. (1975). A class of factor analysis procedures with common asymptotic sampling properties. Psychometrika. 40: 315.

Thangavelu, M.S. and Rajasekharan, S. (1983). Correlation and Path-coefficient analysis in Sesamum. Madras Agric. J. 70: 109-113.

Tikka, S.B.S. and Asawa, B.M. (1978). Factor analysis in lentil. Indian J. Agric. Sci. 48(2): 643-646.

Thurstone, L.L. (1931). Multiple factor analysis. Psychol. Rev. 38: 406-427.

Thurstone, L.L. (1947). Multiple factor analysis. University of Chicago Press, Chicago.

Valsalakumari, P.K., Sivaraman Nair, P.C. and Prabhakaran, P.V. (1985). Genetic divergence in banana. Agric. Res. J. Kerala. 23(2): 146-149.

Wallace, J.T. and Bader, R.S. (1967). Factor analysis in morphometric traits of the house mouse. Syst. Zool. 16(2): 144-148.

Walton, P.D. (1972). Factor analysis of yield in spring wheat. Crop Sci. 12: 731-733.

*Wilks, S.S. (1932 a). Certain generalizations in the analysis of variance. Biometrika 24: 471.

*Wilks, S.S. (1935). On the independence of K sets of normally distributed statistical variables. Econometrics 3: 309.

Wishart, J. (1928). The generalized product moment distribution in sample from a normal multivariate population. Biometrika 20 A: 32.

Yadava, T.P.; Prakash, K. and Yadav, A.K. (1980). Association of yield and its components in sesame. Indian J. Agric. Sci. 4: 317-319.

*Original not seen

APPENDIX

Appendix I Mean values of various characters (Environment 1)

Variety	Characters						
	X_1	X_2	X_3	X_4	X_5	X_6	X_7
1	82.62	2.85	40.08	7.84	6.65	14.50	9.55
2	72.40	2.70	37.68	9.07	3.25	12.30	9.25
3	71.39	2.15	41.71	5.31	1.70	7.00	9.45
4	64.69	1.75	37.17	8.65	2.60	11.25	9.15
5	80.88	3.45	43.35	7.76	12.10	20.35	9.30
6	85.20	2.80	43.49	7.15	5.50	12.65	10.30
7	77.75	2.45	43.61	7.85	5.05	12.90	9.75
8	77.01	2.45	38.62	7.47	4.05	11.50	9.30
9	68.45	1.85	34.85	15.53	3.40	18.90	10.25
10	59.87	2.25	36.91	8.16	4.90	13.05	8.65
11	59.41	2.30	36.28	8.93	2.65	11.60	9.85
12	76.53	2.20	39.38	10.66	4.85	15.50	9.60
13	75.22	2.30	43.01	8.35	4.45	12.80	10.45
14	73.77	2.50	38.27	11.59	4.70	16.20	10.25
15	74.31	1.25	34.05	13.41	1.85	15.25	10.50
16	87.78	1.55	40.99	16.15	4.20	20.35	10.80
17	68.48	1.95	36.61	7.44	1.75	9.20	9.75
18	96.05	1.20	35.02	15.06	6.10	21.15	9.45

Variety	Characters							
	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
1	2.25	2.50	60	43.43	82.5	2.76	46.22	2.18
2	2.10	2.50	58	40.81	82.0	2.86	44.15	1.24
3	2.35	2.45	60	49.07	80.5	3.04	44.02	1.01
4	2.45	2.40	60	41.55	80.5	2.57	44.19	1.34
5	2.25	2.50	54	45.86	81.0	2.89	44.30	2.53
6	2.55	2.65	56	45.28	86.5	2.74	46.68	1.74
7	2.30	2.25	62	42.28	83.5	2.90	47.14	1.92
8	2.40	2.45	67	44.96	82.5	3.05	51.49	1.72
9	1.95	1.80	40	42.72	79.5	2.98	48.15	2.14
10	2.30	2.60	66	48.67	84.5	3.22	49.11	1.79
11	2.55	2.40	72	38.76	84.5	3.03	42.57	1.98
12	2.35	2.50	50	43.10	81.0	2.53	44.30	2.35
13	2.40	2.95	84	47.56	83.5	2.76	48.91	3.17
14	2.25	2.10	46	40.58	78.0	2.80	44.13	1.84
15	2.70	2.60	60	36.45	79.5	2.49	43.60	1.71
16	2.05	2.40	54	35.42	77.0	3.82	45.83	2.69
17	2.40	2.10	50	37.07	82.0	2.98	44.29	1.12
18	2.65	2.75	64	38.94	82.0	2.63	47.13	3.61

Variety	Characters						
	X_1	X_2	X_3	X_4	X_5	X_6	X_7
19	83.83	1.60	44.45	17.13	3.10	20.20	10.95
20	59.15	2.20	35.31	7.15	1.20	8.35	9.30
21	71.36	2.05	33.86	16.83	3.05	19.40	10.35
22	93.69	2.05	41.72	16.31	1.20	28.30	10.35
23	72.57	2.50	29.84	11.90	3.60	15.50	8.50
24	83.67	2.70	41.05	11.54	5.45	17.00	10.05
25	73.22	1.85	33.03	10.65	6.85	17.50	8.10
26	101.24	1.70	42.52	17.19	5.90	23.10	10.35
27	62.13	1.85	35.70	10.24	2.15	12.40	9.35
28	74.65	2.65	31.56	13.16	10.70	18.85	10.30
29	59.04	3.10	32.68	8.02	5.70	13.70	9.35
30	79.56	2.95	39.49	11.00	9.75	20.75	9.95
31	71.33	2.50	34.12	10.99	5.85	16.85	9.25
32	71.56	3.30	35.53	10.67	6.00	16.65	10.20
33	76.50	0.80	37.91	12.36	2.30	14.90	10.60
34	71.00	1.45	29.86	11.05	4.60	15.65	8.50
35	101.59	1.90	40.74	15.51	10.30	25.80	14.95
36	59.61	0.25	29.38	9.20	0.70	9.90	9.55

Variety	Characters							
	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
19	2.35	2.35	62	38.91	83.0	2.58	45.70	2.93
20	2.25	2.20	48	47.46	86.0	2.66	45.51	1.14
21	2.05	2.70	56	43.74	82.5	2.73	45.98	2.22
22	2.50	3.10	70	43.27	82.5	2.71	46.86	3.88
23	2.25	3.10	91	41.63	81.0	2.85	43.38	2.53
24	2.10	2.45	48	40.55	79.5	2.29	46.24	2.23
25	2.40	2.30	52	34.07	75.0	3.04	41.76	1.85
26	2.45	2.50	60	36.89	81.0	2.82	45.74	1.54
27	2.00	2.25	60	43.79	88.0	2.80	45.85	1.66
28	2.15	2.40	54	44.37	86.0	2.42	43.99	2.92
29	2.00	2.10	50	45.93	81.0	2.43	44.31	2.04
30	2.25	2.25	52	46.08	84.5	2.33	45.17	3.14
31	1.95	2.35	50	49.55	82.0	2.34	44.12	3.66
32	2.25	2.25	67	44.28	84.0	2.31	48.40	2.62
33	2.45	3.05	84	38.64	79.5	2.79	46.67	2.15
34	2.65	2.40	50	38.97	85.0	3.18	47.19	1.52
35	2.60	2.75	70	37.84	83.5	2.94	50.20	2.91
36	2.40	2.40	66	44.45	87.5	2.56	49.24	1.38

Variety	Characters						
	X_1	X_2	X_3	X_4	X_5	X_6	X_7
37	73.51	2.6	40.56	10.90	5.75	16.60	9.70
38	80.62	1.85	41.57	12.92	1.70	14.60	9.80
39	93.76	2.75	48.49	12.03	7.10	19.10	10.10
40	88.93	1.70	42.20	15.75	2.85	18.60	9.80
41	72.19	0.70	34.45	10.14	0.20	10.35	9.00
42	77.16	1.70	43.45	11.42	5.50	16.90	9.50
43	79.10	1.85	43.98	13.21	2.35	15.55	9.30
44	109.30	1.90	44.19	15.45	10.85	26.60	9.80
45	69.75	1.90	37.22	12.36	1.40	13.75	9.35
46	81.06	0.35	32.51	17.10	0.05	17.15	10.30
47	66.16	3.25	45.68	3.00	2.60	5.60	7.65
48	76.88	2.20	38.99	11.12	6.75	17.85	9.80
49	79.46	1.75	34.52	12.63	3.60	16.20	8.95
50	80.13	3.05	45.73	7.81	7.50	15.30	8.90
51	75.12	2.20	41.78	9.97	6.25	16.25	9.10
52	69.05	2.50	46.78	6.60	2.80	9.40	8.30
53	75.69	3.15	45.46	7.79	5.60	13.40	9.95
54	74.44	2.65	39.27	7.83	4.35	12.20	9.10

Variety	Characters							
	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
37	2.00	2.25	50	46.11	84.5	2.81	42.49	2.02
38	2.30	3.00	68	47.03	80.5	2.46	43.14	2.98
39	1.95	2.20	46	47.99	92.0	2.40	42.70	1.80
40	2.50	3.05	92	44.95	84.5	2.52	46.46	2.80
41	2.45	3.50	108	44.14	81.0	3.49	44.09	1.18
42	2.50	2.35	49.5	42.73	83.0	3.12	45.37	2.88
43	2.00	2.40	48	38.39	80.5	2.68	44.38	1.65
44	2.50	2.45	74	37.41	87.0	2.70	48.25	3.25
45	2.30	2.50	48	41.43	78.5	3.04	44.87	1.88
46	2.75	2.50	76	39.40	85.5	3.16	47.30	2.22
47	2.00	2.35	42	50.50	79.5	4.12	45.21	0.75
48	2.15	2.50	52	37.27	79.5	3.75	43.55	2.63
49	2.95	2.80	72	41.09	84.0	3.00	47.42	2.42
50	2.10	2.20	50	47.74	81.0	2.81	45.13	1.39
51	2.45	2.50	64	41.47	79.5	2.81	41.6	3.79
52	2.10	2.20	50	45.85	82.5	3.30	42.48	1.04
53	2.30	2.75	52	45.56	20.5	3.23	44.95	3.05
54	2.35	2.30	56	40.24	85.0	2.71	44.06	1.58

Variety	Characters						
	X_1	X_2	X_3	X_4	X_5	X_6	X_7
55	71.13	3.55	42.50	9.80	8.25	18.05	9.15
56	81.35	2.20	44.89	10.49	3.25	13.75	9.25
57	68.95	2.85	35.57	7.38	2.90	10.30	9.95
58	90.61	2.30	53.67	6.35	4.75	11.10	9.40
59	88.35	0.75	34.70	14.86	2.45	17.30	11.05
60	75.97	2.40	47.41	7.09	5.15	12.25	9.90
61	88.00	3.45	57.00	5.75	5.80	11.25	8.70
62	81.03	2.90	52.00	4.18	3.45	7.65	9.20
63	78.82	2.85	42.43	7.87	6.25	14.15	10.55
64	82.77	3.05	44.83	6.31	7.45	13.80	8.70
65	89.16	3.00	52.82	10.88	5.95	16.85	10.45
66	91.88	2.15	35.36	12.66	5.05	17.75	9.05
67	122.58	2.30	47.93	19.01	12.85	31.90	9.15
68	75.10	3.00	41.44	9.33	8.25	17.60	9.90
69	98.53	2.85	58.17	10.99	4.80	15.75	10.10
70	73.75	2.30	48.08	8.27	3.05	11.35	9.50
71	89.76	2.15	50.81	8.17	5.35	13.55	10.15
72	64.08	3.30	40.12	4.25	1.65	5.90	10.25

Variety	Characters							
	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
55	2.40	2.10	54	46.30	79.5	2.55	40.48	3.41
56	2.85	2.60	70	44.27	87.5	2.78	46.11	1.86
57	1.90	2.15	46	42.93	79.0	2.92	47.22	0.78
58	2.45	2.60	58	55.95	86.5	2.69	44.82	1.06
59	2.15	2.55	64	39.61	81.5	2.76	44.13	2.36
60	2.50	2.50	52	45.96	82.0	3.11	40.84	0.81
61	1.85	2.10	50	47.42	85.0	2.99	45.83	1.61
62	2.00	2.15	52	48.35	86.0	3.13	46.56	1.13
63	2.40	2.40	50	46.56	81.0	2.73	46.88	2.29
64	2.45	2.45	56	43.33	82.0	3.23	46.80	2.25
65	2.15	2.35	60	51.05	86.0	3.22	42.51	2.64
66	2.00	2.50	56	39.32	85.5	2.74	42.05	1.85
67	2.85	2.40	62	36.42	84.0	2.54	46.70	5.74
68	2.20	2.50	54	41.30	82.0	2.77	46.25	1.42
69	2.55	2.35	66	42.21	83.0	2.51	47.71	2.29
70	1.95	2.45	42	49.26	83.0	3.07	46.07	0.91
71	2.30	2.55	60	49.67	86.0	2.78	45.66	1.67
72	2.25	2.25	52	49.00	81.5	3.22	49.59	1.21

Variety	Characters						
	X_1	X_2	X_3	X_4	X_5	X_6	X_7
73	78.07	3.35	53.89	4.09	3.15	7.25	9.45
74	73.17	3.15	44.40	7.33	4.40	11.75	8.30
75	87.67	3.30	38.48	7.34	5.95	12.80	9.55
76	74.48	2.85	39.82	7.33	4.00	11.35	9.75
77	81.38	2.10	50.95	9.28	3.50	12.80	10.00
78	78.75	2.85	43.75	12.05	7.15	19.20	10.60
79	84.68	0.60	30.78	14.51	3.10	16.60	9.30
80	81.85	2.50	48.49	9.59	8.85	18.45	9.85
81	110.39	2.25	46.03	17.60	13.60	31.35	9.10
82	85.02	3.25	54.23	8.73	8.95	18.00	9.50
83	80.21	2.20	47.76	9.72	4.00	13.70	9.35
84	73.41	3.75	38.32	7.21	5.65	12.85	9.25
85	101.35	2.30	57.45	9.68	3.95	13.60	10.05
86	74.87	2.40	41.64	8.31	5.20	13.50	9.20
87	77.97	1.85	47.22	4.61	1.25	5.85	9.60
88	76.33	2.10	44.07	8.73	1.65	10.35	9.90
89	94.02	2.40	55.40	7.44	6.80	14.20	9.95
90	95.64	3.10	44.51	9.02	9.10	18.10	10.10

Variety	Characters							
	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
73	2.50	2.40	52	52.06	84.5	2.65	48.06	0.89
74	1.95	2.25	46	43.14	79.5	2.59	46.78	2.49
75	2.30	2.55	54	44.46	79.0	3.41	46.14	1.74
76	2.10	2.20	52	44.12	80.0	3.30	46.52	1.28
77	2.25	2.50	50	47.13	82.0	3.04	44.83	1.44
78	2.45	2.35	58	49.25	86.5	2.41	45.23	1.94
79	2.50	2.50	62	36.96	84.0	2.93	44.24	2.57
80	2.40	2.80	58	45.62	81.5	3.11	44.35	2.35
81	2.25	2.45	62	40.10	79.0	2.40	46.85	4.45
82	2.35	2.55	58	53.59	85.0	2.66	44.08	2.09
83	2.35	2.35	54	47.89	87.0	2.92	45.15	0.96
84	2.00	2.35	56	43.87	84.0	2.82	46.32	2.05
85	2.65	2.55	54	42.34	85.0	2.98	46.68	1.94
86	2.50	2.35	62	46.01	82.5	2.68	46.81	1.75
87	2.45	2.95	102	49.46	84.5	2.43	44.37	1.44
88	2.30	3.00	82	44.93	79.0	2.74	43.72	1.09
89	2.55	2.55	63	48.80	85.5	3.36	46.23	2.07
90	2.80	2.60	70	40.65	84.0	2.44	45.74	2.59

Variety	Characters						
	X_1	X_2	X_3	X_4	X_5	X_6	X_7
91	118.94	2.40	49.06	18.58	18.20	36.80	9.20
92	58.17	2.10	32.02	9.46	1.05	10.50	8.65
93	85.20	2.80	46.34	10.15	9.15	19.30	10.30
94	121.70	1.65	54.95	20.04	7.95	28.00	10.55
95	60.75	2.00	37.23	8.26	2.30	10.55	9.70
96	92.87	1.95	42.92	12.24	7.45	19.70	9.35
97	78.71	2.85	46.00	5.09	3.05	8.15	9.45
98	74.63	2.70	39.30	8.66	7.35	16.00	9.35
99	87.47	2.30	49.68	20.82	7.15	27.95	10.55
100	105.69	2.65	39.69	15.00	8.70	23.7	8.55

Variety	Characters							
	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
91	3.00	2.70	62	37.96	79.5	2.34	45.79	1.43
92	1.90	2.75	83	46.69	85.5	2.47	43.22	1.33
93	2.50	2.60	54	37.85	78.5	2.66	47.23	4.17
94	2.80	2.65	62	38.03	85.0	3.05	47.40	4.52
95	1.85	2.50	62	49.05	81.5	2.47	40.77	1.14
96	2.25	2.60	50	43.61	83.0	2.60	44.20	3.61
97	2.25	2.50	50	47.97	83.0	2.92	44.26	0.79
98	2.05	2.50	43	45.04	80.5	3.11	46.60	3.47
99	2.30	2.40	62	43.41	81.5	2.35	46.72	1.56
100	3.30	2.85	72	40.46	80.5	2.98	44.08	4.16

Appendix II Between dispersion matrix of data A

127.186	-0.543	29.353	16.665	18.348	35.868	1.055	1.378	0.676	12.174	-12.473	3.143	-0.452	4.716	4.322
	0.405	1.666	-1.410	0.626	-0.798	-0.115	-0.059	-0.072	-3.643	1.478	-0.109	-0.005	-0.042	-0.061
		35.010	-7.28	3.271	-3.406	-0.225	0.064	-0.100	-15.103	13.081	3.478	0.169	0.711	-0.713
			9.927	1.826	11.767	0.491	0.268	0.199	8.764	-9.664	-0.586	-0.335	0.613	1.27
				5.721	7.723	-0.036	0.160	-0.031	-6.439	-1.378	0.048	-0.224	0.522	0.876
					19.685	0.388	0.445	0.168	3.035	-10.972	-0.730	-0.544	1.22	2.182
						0.130	0.030	0.018	1.443	-0.562	0.128	-0.041	0.453	0.011
							0.034	0.015	0.932	-0.292	-0.007	-0.005	0.111	0.066
								0.046	2.048	-0.04	-0.067	-0.001	-0.001	0.042
									95.050	-3.530	1.911	-0.638	1.914	1.169
										15.971	3.937	-0.008	-0.533	-1.328
											4.267	-0.221	0.82	-0.248
												0.106	0.029	-0.067
													4.206	0.199
														0.694

Appendix III Within dispersion matrix of data A

94.286	0.722	35.347	19.639	15.714	34.316	3.248	0.654	0.108	7.001	-7.665	0.652	-0.149	-0.101	4.036
	0.169	0.632	0.189	0.355	0.565	0.026	0.016	0.012	0.349	-0.255	0.097	-0.003	-0.003	0.057
		24.044	5.628	5.514	11.238	1.462	0.211	-0.036	3.201	-3.373	0.126	-0.112	-0.046	1.316
			8.992	4.584	13.219	1.179	0.113	0.016	0.813	-0.797	-0.182	-0.011	-0.035	1.508
				8.566	12.445	0.585	0.121	0.039	1.988	-1.422	-0.853	-0.029	-0.012	1.559
					25.130	1.841	0.223	0.068	2.228	-2.223	-0.991	-0.040	-0.053	2.933
						1.115	-0.016	0.004	-1.118	-0.123	0.103	0.01	-0.007	0.214
							0.078	0.027	0.806	-0.155	0.132	0.001	0.001	0.029
								0.047	0.739	-0.038	0.056	0.005	0.002	0.014
									112.258	-1.445	5.958	0.225	-0.099	0.634
										6.517	0.737	-0.008	-0.	-0.319
											7.127	0.059	-0.024	0.008
												0.013	0.002	-0.001
													0.016	-0.002
														0.435

Appendix XV Mean values of various characters (Environment 2)

Variety	Characters						
	X_1	X_2	X_3	X_4	X_5	X_6	X_7
1	92.22	2.57	53.3	13.6	13.7	27.3	9.8
2	86.82	3.40	33.9	17.9	21	38.9	8.2
3	85.24	3.06	48.6	11.7	12.4	24.1	8.6
4	83.30	3.62	50.1	12.1	14.6	26.7	8.4
5	84.93	3.40	48.2	9.5	13.8	23.3	7.8
6	80.37	2.6	47.3	8.3	8.7	17	9.1
7	85.85	2.63	52.3	12.8	16.4	29.2	9.8
8	90.10	3.26	55.5	9.7	9.9	19.6	9.3
9	71.41	2.19	26.7	22.5	8	30.5	8.6
10	78.78	2.79	33	12.4	13.7	26.1	8
11	60.42	1.78	30.7	9.9	4.5	14.4	9.9
12	66.81	1.81	42.6	7.5	5.9	13.4	8.8
13	77.03	3.47	49	7.9	6.5	14.4	8.9
14	74.8	1.93	30.7	16.2	6.9	23.1	8.5
15	85.52	1.21	34.5	21	2.2	23.2	9.7
16	74.46	1.91	27.9	13.3	8.1	21.4	7.8
17	70.04	3.24	30.9	16	15.5	31.5	9.7
18	80.00	1.37	30.5	15.2	8.1	22.8	10.6

	Characters							
Variety	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
1	2.45	2.46	67.2	41.1	84.4	3.16	47.31	3.39
2	2.37	2.28	43.8	40.1	80.8	2.98	44.41	4.22
3	2.20	2.46	46.9	39.5	80.3	3.92	44.00	3.17
4	2.52	2.47	56.4	41.8	81.5	3.01	44.51	3.07
5	2.49	2.42	56.0	40.4	77.9	3.33	44.81	2.97
6	2.55	2.60	54.4	43.8	84.7	3.50	47.61	2.33
7	2.44	2.41	54.6	40.1	82.7	3.36	48.26	4.36
8	2.46	2.47	53.2	38.3	84.5	3.42	51.61	3.04
9	2.46	2.48	50.6	38.1	77.2	2.95	48.61	3.45
10	2.49	2.39	58.8	40.5	81.3	3.40	49.86	3.67
11	2.26	2.45	49.8	34.6	78.6	3.21	43.25	2.48
12	2.79	2.46	59.8	40.9	77.5	3.33	45.55	2.01
13	2.30	3.39	93.2	45.3	88.5	3.59	52.0	3.65
14	2.37	2.16	51.2	38.2	84.3	3.10	44.45	2.68
15	2.55	2.55	49.8	38.4	82.6	2.99	43.85	2.68
16	2.44	2.55	52.6	36.7	75.0	3.22	46.05	3.10
17	2.79	2.40	62.0	37.5	76.2	3.03	45.00	3.62
18	3.00	2.60	83.6	38.3	79.8	3.29	47.30	3.98

Variety	Characters						
	X_1	X_2	X_3	X_4	X_5	X_6	X_7
19	70.40	2.00	31.7	13.4	6.8	20.2	8.7
20	85.37	0.60	25.8	28.1	1.6	29.7	8.6
21	78.37	2.42	44.5	15.5	7.2	22.7	8.7
22	82.07	0.85	36.9	15.9	4.2	20.1	9.5
23	72.09	2.71	29.5	11.4	3.6	15.0	7.0
24	69.65	2.37	28.8	14.3	5.4	19.7	8.6
25	81.98	2.44	27.2	17.7	23.2	40.9	10.6
26	85.51	2.05	32.4	20.3	20.1	40.4	10.7
27	68.60	2.08	29.9	15.0	8.5	23.5	8.4
28	74.05	3.31	38.4	10.6	11.2	21.8	7.4
29	80.56	3.24	40.2	14.2	14.5	28.7	9.0
30	82.43	4.03	36.75	13.7	22.8	36.5	8.2
31	68.48	2.03	36.7	11.5	7.0	18.5	8.8
32	75.88	3.46	35.1	10.8	11.0	21.8	8.0
33	76.05	1.72	28.7	12.4	3.7	15.8	9.9
34	84.16	1.98	46.9	17.2	8.8	26.0	9.8
35	82.49	2.15	37.1	13.6	18.6	32.2	10.7
36	65.62	1.45	32.1	9.4	2.2	11.6	10.4

Variety	Characters							
	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
19	2.34	2.28	49.2	36	78.0	2.85	46.15	2.94
20	2.14	2.5	55.8	37.4	81.8	3.32	46.60	3.46
21	2.37	2.36	47.7	40.8	80.4	3.25	46.05	3.38
22	2.64	2.53	54.2	36.8	86.4	3.17	47.65	2.89
23	2.02	3.22	108.3	38.5	76.7	3.37	47.60	2.42
24	2.24	2.24	46.2	39.5	78.2	2.96	47.10	2.56
25	2.57	2.49	60.2	37.9	87.8	3.10	42.85	4.84
26	2.52	2.54	57.6	38.2	78.5	3.55	46.31	3.04
27	2.16	2.28	47.2	37.4	75.3	3.05	46.11	3.16
28	2.18	2.16	49.4	40.1	79.3	3.06	44.51	2.66
29	2.46	2.4	56.8	40.0	82.5	3.06	44.60	4.25
30	2.36	2.32	52.4	41.2	81.9	3.03	48.26	5.57
31	2.04	2.29	47.6	41.3	83.9	3.05	44.60	4.02
32	2.35	2.41	53.4	41.1	85.9	2.84	48.85	3.45
33	2.59	3.29	105.4	37.1	79.0	3.10	49.65	2.30
34	2.05	2.42	44.4	39.6	82.2	3.30	48.40	2.51
35	2.59	2.75	66.0	41.4	80.6	3.60	52.80	3.61
36	2.53	3.03	77.8	41.8	77.7	2.7	50.50	1.61

	Characters						
Variety	X_1	X_2	X_3	X_4	X_5	X_6	X_7
37	80.61	5.58	32.2	13.8	14.5	28.3	8.0
38	62.36	1.32	34.1	9.0	2.7	11.7	9.7
39	94.27	3.35	53.9	12.0	8.8	20.8	10.2
40	69.94	0.94	31.3	10.7	1.5	12.2	11.3
41	69.74	2.24	26.3	12.3	6.4	18.7	8.1
42	74.14	1.27	34.4	12.2	2.1	14.3	10.5
43	70.66	2.23	30.7	15.1	7.0	22.1	8.2
44	79.32	1.79	41.8	11.3	6.0	17.3	12.2
45	69.95	2.67	28.6	15.0	17.1	32.1	8.5
46	86.68	0.97	37.2	18.8	10.7	29.5	10.2
47	92.27	3.70	61.9	10.6	14.1	24.7	9.1
48	72.92	3.33	39.9	10.1	15.1	25.2	9.0
49	70.83	1.86	32.9	9.3	4.2	13.5	7.7
50	76.20	4.36	39.7	15.6	22.8	38.4	11.2
51	83.86	2.66	40.0	15.5	13.3	28.5	11.2
52	79.46	2.39	47.1	12.0	8.5	20.5	9.4
53	75.88	2.66	43.4	10.5	4.9	15.4	9.1
54	81.55	2.92	47.8	12.4	15.4	27.8	9.1

Variety	Characters							
	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
37	2.46	2.37	54.4	39.4	79.7	2.85	42.81	3.30
38	2.52	2.73	58.2	40.7	76.6	2.64	43.85	2.28
39	1.92	2.45	46.2	47.9	92.1	2.38	43.5	1.96
40	2.35	2.93	69.4	42.3	81.5	2.65	47.1	1.65
41	2.53	3.35	92.8	37.8	82.9	4.21	49.65	2.08
42	2.64	2.49	62.0	41.9	82.7	3.59	46.15	3.45
43	2.40	2.27	42.9	39.5	75.9	3.41	48.75	2.34
44	2.83	2.50	68.2	40.9	89.0	3.09	48.5	2.85
45	2.40	2.35	47.5	38.3	79.2	3.36	48.6	4.38
46	3.01	2.88	78.2	38.8	82.5	3.82	48.7	4.04
47	2.47	2.44	50.6	41.9	91.5	4.56	43.4	3.12
48	2.77	2.72	55.4	38.7	81.2	4.26	44.05	3.77
49	2.63	2.55	55.2	37.9	79.6	3.12	48.05	2.02
50	2.27	2.45	53.0	43.8	82.1	3.21	45.75	3.62
51	2.53	2.56	58.4	39.6	79.8	2.59	42.54	6.42
52	2.63	2.47	55.0	42.7	88.2	3.40	43.04	2.32
53	2.36	2.63	52.2	39.9	81.3	3.39	46.39	3.50
54	2.45	2.46	57.8	43.0	84.4	2.77	44.24	3.41

Variety	Characters						
	X_1	X_2	X_3	X_4	X_5	X_6	X_7
55	74.67	2.79	40.4	10.6	7.2	17.8	8.1
56	81.01	2.09	46.6	9.6	4.2	13.8	9.0
57	74.09	3.02	35.7	12.6	9.8	22.4	8.4
58	89.35	2.96	48.7	12.1	8.4	20.5	9.75
59	76.85	1.08	32.8	14.7	3.5	18.2	9.0
60	64.82	1.78	48.2	6.8	3.9	10.7	9.1
61	83.58	2.81	55.5	12.1	11.5	23.6	8.5
62	108.98	5.34	74.6	12.5	22.5	35.0	11.6
63	68.90	2.50	42.1	9.1	5.5	14.6	9.3
64	90.96	3.06	45.6	11.4	14.8	26.2	7.7
65	79.79	2.43	42.0	13.2	9.9	23.1	8.1
66	89.62	1.94	40.1	18.4	3.7	22.1	7.4
67	90.01	3.27	41.8	17.1	9.4	26.5	9.2
68	76.76	3.7	39.8	11.6	15.7	27.3	9.1
69	76.17	2.23	50.4	8.3	5.2	13.5	8.3
70	75.24	1.62	45.1	11.3	4.3	15.6	9.9
71	78.09	2.98	43.7	11.4	8.1	19.5	8.7
72	80.58	2.81	40.9	14.1	15.0	29.0	9.4
73	89.21	3.37	53.1	11.4	12.5	23.9	8.9

Variety	Characters							
	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
55	2.70	2.51	55.8	40.8	86.1	2.89	40.24	3.10
56	2.78	2.58	56.3	39.4	85.5	3.25	47.3	2.71
57	2.52	2.63	70.8	37.2	70.8	2.33	48.75	1.65
58	2.40	2.42	52.8	47.2	89.5	2.92	45.1	2.08
59	2.62	2.65	59.2	37.6	81.4	2.75	44.79	2.58
60	2.55	2.55	56.8	40.9	87.1	3.40	40.75	1.29
61	2.30	2.38	58.8	44.6	90.4	3.41	46.79	3.58
62	2.43	2.44	58.2	49.8	88.0	3.52	46.64	5.08
63	2.50	2.79	49.3	42.3	91.1	3.70	48.09	2.32
64	2.44	2.31	52.0	41.8	86.7	3.09	47.59	3.35
65	2.24	2.33	40.8	45.7	84.6	3.04	46.24	3.68
66	2.10	2.50	54.3	40.6	79.6	2.69	42.75	2.32
67	2.96	2.50	71.6	41.3	80.6	2.77	47.25	4.78
68	2.47	2.56	57.6	41.2	77.0	2.78	47.80	2.38
69	3.41	3.03	65.2	43.1	77.8	2.93	48.45	1.95
70	2.49	2.44	53.4	46.6	90.2	3.08	46.60	1.23
71	2.35	2.42	51.8	46.5	91.6	3.40	45.79	2.40
72	2.57	2.43	59.4	40.5	84.2	3.56	50.84	3.75
73	2.59	2.47	61.6	46.8	86.8	2.89	48.69	2.72

Variety	Characters						
	X_1	X_2	X_3	X_4	X_5	X_6	X_7
74	77.97	2.43	43.9	9.8	8.5	18.3	7.6
75	83.59	2.80	44.8	10.0	13.4	23.4	7.5
76	76.53	2.91	42.5	9.7	10.9	20.6	7.7
77	85.31	2.64	48.1	12.3	8.6	20.9	8.9
78	80.37	2.07	45.0	13.0	9.0	22.0	8.8
79	79.68	2.10	33.1	22.8	14.1	36.9	9.0
80	91.84	2.49	50.6	11.6	12.5	24.1	7.2
81	75.63	1.81	40.7	11.7	5.1	16.8	10.1
82	83.22	2.34	44.4	11.5	9.9	21.4	8.9
83	74.35	2.60	45.6	11.6	5.6	17.2	9.2
84	77.41	3.66	43.5	8.8	11.2	20.0	9.1
85	84.23	2.84	51.6	11.5	3.2	14.7	11.0
86	80.87	3.14	47.7	10.8	12.3	22.6	8.6
87	74.76	1.17	49.0	6.7	1.6	8.3	8.5
88	90.01	3.40	32.8	18.0	29.6	47.6	8.4
89	87.42	3.93	58.3	9.4	13.8	23.2	8.8
90	90.19	3.13	46.8	10.3	14.2	24.5	8.3

Variety	Characters							
	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
74	2.27	2.41	49.6	44.1	86.6	3.35	46.59	3.88
75	2.44	2.47	60.0	43.0	90.6	3.42	47.29	3.00
76	2.30	2.37	58.6	40.2	87.9	3.08	47.85	2.97
77	2.53	2.45	58.9	42.3	87.4	3.08	45.65	2.38
78	2.54	2.37	51.6	44.0	88.2	3.00	46.60	2.18
79	2.75	2.51	60.6	37.8	81.0	3.00	43.10	3.64
80	2.35	2.46	52.6	42.4	84.2	3.14	44.50	3.06
81	2.75	2.44	65.8	42.2	80.9	2.79	47.80	2.38
82	2.42	2.38	55.4	44.2	80.0	3.24	45.20	2.55
83	2.34	3.59	96.6	46.0	82.7	3.45	45.70	2.01
84	2.40	2.46	53.4	43.0	88.1	3.67	46.60	3.21
85	2.77	2.51	67.1	47.3	88.1	3.31	47.50	2.00
86	2.54	2.52	61.0	40.7	78.7	2.88	48.40	3.13
87	2.41	3.59	96.2	46.6	89.1	2.99	45.30	1.97
88	2.44	2.43	52.0	38.7	84.6	2.93	44.60	4.84
89	2.36	2.46	59.4	44.0	88.8	3.25	45.60	3.29
90	2.65	2.47	58.8	39.1	90.0	3.35	46.10	3.37

Variety	Characters						
	X_1	X_2	X_3	X_4	X_5	X_6	X_7
91	71.21	2.19	30.5	11.2	11.4	22.60	8.3
92	70.60	0.72	32.4	14.8	3.2	18.00	10.7
93	76.43	2.88	35.1	11.4	12.9	24.30	8.1
94	85.89	1.84	37.7	14.8	11.9	26.20	11.5
95	74.52	3.21	31.2	15.5	16.2	31.70	9.2
96	74.95	2.02	37.1	11.0	6.1	17.10	8.2
97	87.74	2.55	59.1	11.5	8.3	19.30	8.3
98	73.29	0.01	27.8	12.95	2.0	13.15	8.7
99	85.20	2.81	35.7	17.6	13.4	31.10	9.1
100	70.27	1.20	35.7	7.8	2.1	9.90	7.2

Variety	Characters							
	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
91	2.83	2.24	59.4	37.5	78.6	2.85	43.6	3.59
92	2.34	2.42	56.2	39.6	78.4	2.70	40.45	2.33
93	2.42	2.37	52.6	40.0	86.7	3.25	43.79	5.93
94	2.78	2.43	57.8	40.4	85.3	3.12	47.85	3.49
95	2.39	2.32	50.8	40.7	87.5	3.05	41.35	3.54
96	2.32	2.45	49.6	41.7	84.4	2.95	44.25	3.05
97	2.45	2.35	54.4	44.5	83.4	3.39	44.55	2.08
98	2.47	2.38	58.4	36.9	79.8	3.39	44.65	4.39
99	2.75	2.70	59.0	39.9	77.9	2.77	47.20	3.27
100	3.23	2.65	69.9	41.5	83.4	3.17	44.60	1.70

Appendix V Between dispersion matrix of data B

35.025	2.164	33.708	-0.737	10.435	9.685	0.088	-0.212	-0.480	-13.184	8.878	12.489	0.119	1.168	0.492
	0.536	2.993	-0.686	2.571	1.893	-0.197	-0.038	-0.060	-1.820	0.736	0.586	0.043	-0.007	0.156
		67.092	-14.849	4.589	-10.219	0.190	-0.075	-0.074	-7.101	18.487	21.002	0.599	0.428	-1.021
			8.188	1.067	9.251	0.157	-0.124	-0.260	-8.482	-3.784	-4.144	-0.222	-0.633	0.417
				18.972	20.074	-0.155	-0.129	-0.613	-19.937	0.075	-0.097	0.211	-0.134	1.475
					29.354	-0.02	-0.257	-0.873	-28.655	-3.715	-4.184	-0.007	-0.787	1.875
						0.837	0.044	0.034	1.785	0.298	0.331	-0.002	0.228	0.013
							0.049	0.009	0.655	-0.076	-0.104	-0.003	0.064	-0.026
								0.061	2.702	0.116	-0.008	0.017	0.193	-0.084
									140.524	1.63	-4.581	0.368	9.832	-2.637
										7.339	7.35	0.008	0.305	-0.588
											16.322	0.371	-0.821	-0.003
												0.087	0.172	0.022
													5.93	-0.100
														0.251

Appendix VI Within dispersion matrix of data B

56.188	1.882	21.908	14.441	21.499	35.779	1.899	0.279	0.226	6.199	-0.586	3.516	0.261	0.053	4.001
	0.659	1.413	0.347	2.196	2.541	0.091	-0.002	0.008	-0.383	0.271	0.386	-0.019	-0.002	0.264
		32.961	1.691	7.044	8.584	1.233	0.043	0.019	-3.367	0.791	2.695	0.181	0.069	0.797
			9.472	7.170	16.636	0.959	0.109	0.089	1.526	-0.582	-0.135	0.059	-0.007	1.579
				27.569	34.673	0.782	0.125	0.209	3.379	0.786	4.155	-0.062	-0.004	3.708
					51.260	1.736	0.232	0.296	4.887	0.248	3.981	-0.005	-0.011	5.302
						0.627	0.013	-0.009	-0.252	0.254	0.367	0.012	-0.007	0.124
							0.014	0.003	0.275	0.005	0.034	0.004	0.003	0.04
								0.025	0.345	-0.056	-0.067	-0.006	-0.001	0.019
									36.937	-0.428	0.404	-0.194	-0.042	0.952
										2.730	1.048	0.025	-0.005	0.152
											8.172	0.153	0.015	0.073
												0.065	0.001	-0.004
													0.007	0.001
														1.294

APPENDIX-VII

COMPUTER PROGRAMMES(BASIC) USED FOR THE ANALYSIS

```
5 REM PROGRAM FOR COMPUTATION OF PRINCIPAL FACTOR SOLUTION
10 DIM D(20,20),A(20,20),V(20),AJ(20),FMT(20)
20 DIM W(20,20),U(20),B(20,20),EA(20,20),EV(20,20),EEA(20,20),EEV(20,20),EP(20),EMM(20)
25 DIM DIAG(20)
27 DIM VA(20,20),VC(20)
30 INPUT "N,NF,EPS,NO";N,NF,EPS,NO
32 INPUT "FILE NAME";FL$:OPEN "I",#1,FL$
34 INPUT "EIGEN VAL. LIMIT";VCONST
36 INPUT "1st trial [Y/ ]";FT$
37 IF FT$<>"Y" THEN Z=1:INPUT "NO OF FACTORS";M:FOR I=1 TO N:INPUT "DIAG ELTS";DIAG(I):NEXT I
40 IR=0
45 RESTORE
70 FOR I=1 TO N
80 FOR J=I TO N
92 INPUT #1,D(I,J)
100 D(J,I)=D(I,J)
110 NEXT J
120 NEXT I
124 TRES=0
125 IF Z=0 THEN 130
126 FOR I=1 TO N
127 D(I,I)=DIAG(I)
128 NEXT I
130 FOR I=1 TO N
140 FOR J=1 TO N
150 EA(I,J)=D(I,J)
160 NEXT J ':LPRINT
161 NEXT I
162 FOR I=1 TO N
164 TRES=TRES+D(I,I)
166 NEXT I
170 GOSUB 1530
172 FOR I=1 TO N
174 V(I)=EA(I,I)
176 FOR J=1 TO N
177 W(I,J)=EV(I,J)
```

```
178 NEXT J:NEXT I
180 LPRINT "  ESTIMATES OF PRINCIPAL FACTOR LOADINGS"
181 IF Z=1 THEN 190
182 M=N
183 RN=0
184 FOR I=1 TO N
185 IF V(I)<VCONST THEN RN=RN+1
186 NEXT I
187 M=M-RN
190 FOR I=1 TO M
200 SS=0
210 XP=SQR(V(I))
220 FOR J=1 TO N
230 SS=SS+W(J,I)*W(J,I)
240 NEXT J
250 SS=XP/SQR(SS)
260 FOR J=1 TO N
270 A(I,J)=SS*W(J,I)
280 NEXT J:NEXT I
290 FOR J=1 TO N
300 LPRINT J;"      ";
305 DIAG(J)=0
310 FOR I=1 TO M
315 DIAG(J)=DIAG(J)+A(I,J)*A(I,J)
320 LPRINT A(I,J);
330 NEXT I:LPRINT:NEXT J
331 LPRINT:LPRINT "DIAG ELTS":FOR I=1 TO N:LPRINT DIAG(I);:NEXT I:LPRINT
332 Z=1
333 FOR I=1 TO N
334 IF ABS(DIAG(I)-D(I,I))>EPS THEN GOTO 122
335 NEXT I
340 NFF=M
345 LPRINT:LPRINT "VARIMAX ROTATION WITH ESTIMATES OF PRINCIPAL FACTOR LOADINGS"
347 FOR I=1 TO NFF:FOR J=1 TO N:VA(J,I)=A(I,J):NEXT J,I
348 GOSUB 3020
1525 END
1527 REM SUBR.BRANCHING FROM 170
1530 NN=1
1540 IF NN=0 THEN 1620
1550 FOR EI=1 TO N
1560 FOR EJ=1 TO N
1570 IF EI<>EJ THEN 1600
1580 EV(EI,EJ)=1
```

```
1590 GOTO 1610
1600 EV(EI,EJ)=0
1610 NEXT EJ,EI
1620 ENR=0
1630 EMI=N-1
1640 FOR EI=1 TO EMI
1650 EP(EI)=0
1660 EMJ=EI+1
1670 FOR EJ=EMJ TO N
1680 IF EP(EI)>ABS(EA(EI,EJ)) THEN 1710
1690 EP(EI)=ABS(EA(EI,EJ))
1700 EMM(EI)=EJ
1710 NEXT EJ:NEXT EI
1720 FOR EI=1 TO EMI
1730 IF EI<=1 THEN 1750
1740 IF EPMAX>EP(EI) THEN 1780
1750 EPMAX=EP(EI)
1760 EIP=EI
1770 EJP=EMM(EI)
1780 NEXT EI
1790 IF ENR=0 THEN EEPLN=ABS(EPMAX)*9.999999E-10
1800 IF EPMAX<=EEPLN THEN 2641
1810 ENR=ENR+1
1820 IF EA(EIP,EIP))=EA(EJP,EJP) THEN 1860
1830 ETA=-2*EA(EIP,EJP)/(ABS(EA(EIP,EIP)-EA(EJP,EJP))+SQR((EA(EIP,EIP)-EA(EJP,EJP))^2+4*EA(EIP,EJP)^2))
1840 GOTO 1870
1850 GOTO 1870
1860 ETA=2*EA(EIP,EJP)/(ABS(EA(EIP,EIP)-EA(EJP,EJP))+SQR((EA(EIP,EIP)-EA(EJP,EJP))^2+4*EA(EIP,EJP)^2))
1870 ECO=1/SQR((1+ETA*ETA))
1880 ESI=ETA*ECO
1890 EAI=EA(EIP,EIP)
1900 EA(EIP,EIP)=ECO*ECO*(EAI+ETA*(2*EA(EIP,EJP)+ETA*(EA(EJP,EJP))))
1910 EA(EJP,EJP)=ECO*ECO*(EA(EJP,EJP)-ETA*(2*EA(EIP,EJP)-ETA*EAI))
```

```
1920 EA(EIP,EJP)=0
1930 IF EA(EIP,EIP))=EA(EJP,EJP) THEN 2030
1940 ETT=EA(EIP,EIP)
1950 EA(EIP,EIP)=EA(EJP,EJP)
1960 EA(EJP,EJP)=ETT
1970 IF ESI)=0 THEN 2000
1980 ETT=ECO
1990 GOTO 2010
2000 ETT=-ECO
2010 ECO=ABS(ESI)
2020 ESI=ETT
2030 FOR EI=1 TO EMI
2040 IF EI-EIP)0 THEN 2070
2050 IF EI-EIP(0 THEN 2080
2060 IF EI-EIP=0 THEN 2210
2070 IF EI=EJP THEN 2210
2080 IF EMM(EI)=EIP THEN 2100
2090 IF EMM(EI)()EJP THEN 2210
2100 EK=EMM(EI)
2110 ETT=EA(EI,EK)
2120 EA(EI,EK)=0
2130 EMJ=EI+1
2140 EP(EI)=0
2150 FOR EJ=EMJ TO N
2160 IF EP(EI))ABS(EA(EI,EJ)) THEN 2190
2170 EP(EI)=ABS(EA(EI,EJ))
2180 EMM(EI)=EJ
2190 NEXT EJ
2200 EA(EI,EK)=ETT
2210 NEXT EI
2230 EP(EIP)=0
2240 EP(EJP)=0
```

```
2250 FOR EI=1 TO N
2260 IF EI>EIP THEN 2380
2270 IF EI=EIP THEN 2570
2280 ETT=EA(EI,EIP)
2290 EA(EI,EIP)=ECO*ETT+ESI*EA(EI,EJP)
2300 IF EP(EI)>=ABS(EA(EI,EIP)) THEN 2330
2310 EP(EI)=ABS(EA(EI,EIP))
2320 EMM(EI)=EIP
2330 EA(EI,EJP)=-ESI*ETT+ECO*EA(EI,EJP)
2340 IF EP(EI)>=ABS(EA(EI,EJP)) THEN 2570
2350 EP(EI)=ABS(EA(EI,EJP))
2360 EMM(EI)=EJP
2370 GOTO 2570
2380 IF EI>EJP THEN 2480
2390 IF EI=EJP THEN 2570
2400 ETT=EA(EIP,EI)
2410 EA(EIP,EI)=ECO*ETT+ESI*EA(EI,EJP)
2420 IF EP(EIP)>=ABS(EA(EIP,EI)) THEN 2450
2430 EP(EIP)=ABS(EA(EIP,EI))
2440 EMM(EIP)=EI
2450 EA(EI,EJP)=-ETT*ESI+ECO*EA(EI,EJP)
2460 IF EP(EI)>=ABS(EA(EI,EJP)) THEN 2570
2470 GOTO 2350
2480 ETT=EA(EIP,EI)
2490 EA(EIP,EI)=ETT*ECO+ESI*EA(EJP,EI)
2500 IF EP(EIP)>=ABS(EA(EIP,EI)) THEN 2530
2510 EP(EIP)=ABS(EA(EIP,EI))
2520 EMM(EIP)=EI
2530 EA(EJP,EI)=-ETT*ESI+ECO*EA(EJP,EI)
2540 IF EP(EJP)>=ABS(EA(EJP,EI)) THEN 2570
2550 EP(EJP)=ABS(EA(EJP,EI))
2560 EMM(EJP)=EI
2570 NEXT EI
2580 IF NN=0 THEN 1720
2590 FOR EI=1 TO N
2600 ETT=EV(EI,EIP)
```

```
2610 EV(EI,EIP)=ETT*ECO+ESI*EV(EI,EJP)
2620 EV(EI,EJP)=-ETT*ESI+ECO*EV(EI,EJP)
2630 NEXT EI
2640 GOTO 1720
2641 FOR EI=1 TO N
2642 EP(EI)=1:SUM1=0
2646 FOR EJ=1 TO N:SUM1=SUM1+EV(EJ,EI)*EA(EJ,EJ)
2648 NEXT EJ
2649 IF SUM1<0 THEN EP(EI)=-1
2650 NEXT EI
2652 FOR EI=1 TO N
2654 FOR EJ=1 TO N
2656 EV(EJ,EI)=EV(EJ,EI)*EP(EI)
2658 NEXT EJ:NEXT EI
2659 LPRINT "NO OF ROTATIONS";ENR
2660 LPRINT "EIGEN VALUES & CORR. EIGEN VECTORS":FOR EI=1 TO N
2670 REM LPRINT "EIGEN VALUES",EA(EI,EI)
2672 LPRINT EA(EI,EI);"       ";
2680 REM LPRINT "CORRESPONDING EIGEN VECTORS"
2690 FOR EJ=1 TO N
2700 LPRINT USING "##.####";EV(EJ,EI);
2710 NEXT EJ
2720 LPRINT
2730 NEXT EI
2740 RETURN
3010 REM SUBROUTINE BRANCHING FROM 1524
3020 REM INPUT "NO OF FACTORS, NO OF VARIABLES";VN,VM
3022 VN=NFF:VM=N
3030 EP=.00116
3032 GOTO 3080
3040 FOR VJ=1 TO VN
3050 FOR VK=1 TO VM
3060 VA(VK,VJ)=A(VK,VJ)
3062 REM INPUT "VA(k,j)";VA(VK,VJ)
3070 NEXT VK : NEXT VJ
```

```
3080 FOR VJ=1 TO VM
3090 VC(VJ)=0
3100 FOR VK=1 TO VN
3110 VC(VJ)=VC(VJ) +VA(VJ,VK)^2
3120 NEXT VK
3130 LPRINT "VARIABLE";VJ,"COMMUNALITY";
3140 VC(VJ)=SQR(VC(VJ))
3150 FOR VK=1 TO VN
3160 VA(VJ,VK)=VA(VJ,VK)/VC(VJ)
3170 NEXT VK: NEXT VJ
3180 VN1=VN-1
3190 VNR=0
3200 FOR VI=1 TO VN1
3210 VI1=VI+1
3220 FOR VJ=VI1 TO VN
3230 VA1=0
3240 VB1=0
3250 VC1=0
3260 VD1=0
3270 FOR VK=1 TO VM
3280 VU=VA(VK,VI)^2 -VA(VK,VJ)^2
3290 VV=VA(VK,VI)*VA(VK,VJ)*2
3300 VA1=VA1+VU
3310 VB1=VB1+VV
3320 VC1=VC1+VU^2-VV^2
3330 VD1=VD1+VU*VV*2
3340 NEXT VK
3350 VQN =VD1-2*VA1*VB1/VM
3360 VQD=VC1-(VA1^2-VB1^2)/VM
3370 IF ABS(VQN)-ABS(VQD)<0 THEN 3710
3380 IF ABS(VQN)-ABS(VQD)>0 THEN 3490
3390 IF ABS(VQN)-ABS(VQD)=0 THEN 3570
3400 VEM=ABS(VQN/VQD)
3410 IF VEM-EP <0 THEN 3450
```

XXXV

;VC(VJ)

```
3420 VCS=COS(ATAN(VEM))
3430 VSN=SIN(ATAN(VEM))
3440 GOTO 3590
3450 IF VQD>=0 THEN 3710
3460 VSP=.70710678#
3470 VCP=VSP
3480 GOTO 3740
3490 VEM=ABS(VQD/VQN)
3500 IF VEM<VEP THEN 3540
3510 VSN=1!/SQR(1+VEM^2)
3520 VCS=VSN*VEM
3530 GOTO 3590
3540 VCS=0
3550 VCN=1
3560 GOTO 3590
3570 VCS=.70710678#
3580 VSN=VCS
3590 VEM=SQR((1+VCS)*.5)
3600 VCS1=SQR((1+VEM)*.5)
3610 VSN1=VSN/(4*VCS1*VEM)
3620 IF VQD>=0 THEN 3660
3630 VCP=.70710678#*(VCS1+VSN1)
3640 VSP=.70710678#*(VCS1-VSN1)
3650 GOTO 3680
3660 VCP=VCS1
3670 VSP=VSN1
3680 IF VQN>=0 THEN 3730
3690 VSP=-VSP
3700 GOTO 3730
3710 VNR=VNR+1
3720 GOTO 3780
3730 FOR VK=1 TO VM
```

```
3740 VEM=VA(VK,VI)*VCP+VA(VK,VJ)*VSP
3750 VA(VK,VJ)=VA(VK,VJ)*VCP-VA(VK,VI)*VSP
3760 VA(VK,VI)=VEM
3770 NEXT VK
3780 NEXT VJ
3790 NEXT VI
3800 IF VNR<>(VN*VN1)/2 THEN 3190
3810 FOR VK=1 TO VM
3820 FOR VL=1 TO VN
3830 VA(VK,VL)=VA(VK,VL)*VC(VK)
3840 NEXT VL:NEXT VK
3850 LPRINT "NEW FACTOR PATTERN"
3860 FOR VJ=1 TO VM
3870 LPRINT VJ
3880 FOR VK=1 TO VN
3890 LPRINT USING "###.#####";VA(VJ,VK);
3900 NEXT VK:NEXT VJ
3910 FOR VJ=1 TO VN
3920 VC(VJ)=0
3930 FOR VK=1 TO VM
3940 VC(VJ)=VC(VJ)+VA(VK,VJ)^2
3950 NEXT VK:NEXT VJ
3960 FOR VJ=1 TO VN
3970 VC(VJ)=VC(VJ)/VM
3980 NEXT VJ
3985 LPRINT
3990 LPRINT "PROP. VAR ACCOUNTED BY EACH FACTOR"
4000 FOR VJ=1 TO VN
4010 LPRINT VJ,VC(VJ)
4015 NEXT VJ
4020 RETURN
```

```
5 REM PROGRAM FOR COMPUTATION OF MAXIMUM LIKELIHOOD SOLUTION
10 DIM D(20,20),A(20,20),V(20),AJ(20),FMT(20)
20 DIM W(20,20),U(20),B(20,20),EA(20,20),EV(20,20),EEA(20,20),EEV(20,20),EP(20),EMM(20)
25 DIM DIAG(20)
27 DIM VA(20,20),VC(20)
30 INPUT "N,NF,EPS,NO";N,NF,EPS,NO
32 INPUT "FILE NAME";FL$:OPEN "I",#1,FL$
70 FOR I=1 TO N
80 FOR J=1 TO N
92 INPUT #1,D(I,J)
100 D(J,I)=D(I,J)
110 NEXT J
120 NEXT I
130 NFF=NF
140 FOR J=1 TO N
150 FOR I=1 TO NFF
160 PRINT "A";I,J
170 INPUT A(I,J)
180 NEXT I:NEXT J
360 FOR I=1 TO N
370 V(I)=1
380 FOR J=1 TO NFF
390 V(I)=V(I)-A(J,I)*A(J,I)
400 NEXT J:LPRINT V(I);: NEXT I:LPRINT
410 FOR I=1 TO NFF
420 FOR J=1 TO N
430 W(I,J)=A(I,J)/V(J)
440 NEXT J:NEXT I
450 FOR I=1 TO NFF
460 FOR J=1 TO N
470 B(I,J)=A(I,J)
480 NEXT J:NEXT I
510 FOR I=1 TO N
520 U(I)=0
522 FOR J=1 TO N
530 U(I)=U(I)+W(1,J)*D(I,J)
535 NEXT J
540 U(I)=U(I)-A(1,I)
550 NEXT I
560 H=0
570 FOR I=1 TO N
580 H=H+U(I)*W(1,I)
585 NEXT I
590 H=ABS(H)
600 H=1/SQR(H)
610 FOR I=1 TO N
```

```
620 A(1,I)=H*U(I)
630 NEXT I
640 FOR I=2 TO NFF
650 II=I-1
660 FOR J=1 TO II
670 AJ(J)=0
680 FOR K=1 TO N
690 AJ(J)=AJ(J)+A(J,K)*W(I,K)
700 NEXT K:NEXT J
710 FOR KK=1 TO N
720 U(KK)=0
730 FOR J=1 TO N
740 U(KK)=U(KK)+W(I,J)*D(KK,J)
750 NEXT J
760 FOR J=1 TO II
770 U(KK)=U(KK)-AJ(J)*A(J,KK)
780 NEXT J
790 U(KK)=U(KK)-A(I,KK)
800 NEXT KK
810 H=0
820 FOR J=1 TO N
830 H=H+U(J)*W(I,J)
840 NEXT J
850 H=ABS(H)
860 H=1/SQR(H)
870 FOR J=1 TO N
880 A(I,J)=H*U(J)
890 NEXT J
900 NEXT I
910 FOR I=1 TO NFF
920 FOR J=1 TO N
930 DIF=ABS(A(I,J)-B(I,J))
940 IF DIF>EPS THEN 970
950 NEXT J:NEXT I
960 GOTO 1010
970 FOR I=1 TO NFF
980 FOR J=1 TO N
```

```
985 B(I,J)=A(I,J)
990 NEXT J:NEXT I
991 FOR J=1 TO N
992 LPRINT J
993 FOR I=1 TO NFF
994 LPRINT A(I,J);
995 NEXT I:LPRINT:NEXT J
1000 GOTO 360
1010 LPRINT "MAX. LIKELY ESTIMATES OF FACTOR LOADINGS"
1020 FOR J=1 TO N
1030 LPRINT J
1040 FOR I=1 TO NFF
1050 LPRINT A(I,J);
1060 NEXT I:LPRINT:NEXT J
1062 LPRINT:LPRINT "Coefft. of Uniqueness"
1070 FOR I=1 TO N
1080 'V(I)=D(I,I)
1082 V(I)=1
1090 FOR J=1 TO NFF
1100 V(I)=V(I)-A(J,I)*A(J,I)
1110 NEXT J:LPRINT V(I);:NEXT I:LPRINT
1120 LPRINT "FACTOR VALUE","P.C. VARIATION EXPLAINED"
1130 FOR I=1 TO NFF
1140 XX=0
1150 FOR J=1 TO N
1160 XX=XX+A(I,J)*A(I,J)
1170 NEXT J
1180 H=XX/TRES*100
1190 LPRINT I,XX,H
1200 NEXT I
1210 FOR I=1 TO N
1220 B(I,I)=V(I)
1230 NEXT I
1240 FOR I=1 TO NFF
1250 FOR J=1 TO N
1260 W(J,I)=A(I,J)
```

```
1270 NEXT J:NEXT I
1280 FOR I=1 TO N
1282 B(I,I)=0
1290 FOR J=I+1 TO N
1300 IF I=J THEN 1370
1310 B(I,J)=0
1320 FOR K=1 TO NFF
1330 B(I,J)=B(I,J)+A(K,I)*W(J,K)
1340 NEXT K
1350 B(I,J)=D(I,J)-B(I,J)
1360 NEXT J:NEXT I
1370 LPRINT "RESIDUAL MATRIX AFTER REMOVAL OF";
1380 FOR I=1 TO N
1390 LPRINT I
1400 FOR J=1 TO I
1410 LPRINT B(J,I);
1420 NEXT J:LPRINT:NEXT I
1422 LPRINT
1430 XX=0
1440 NCC=N-1
1450 FOR I=1 TO NCC
1460 II=I+1
1470 FOR J=II TO N
1480 XX=XX+B(I,J)*B(I,J)/(V(I)*V(J))
1485 NEXT J:NEXT I
1490 OON=NO-.1666667#*(2*N+5)-.6666667*NFF
1500 DF=.5*((N-NFF)^2-N-NFF)
1510 XX=XX*OON
1520 LPRINT "APP. CHI-SQ";XX;" WITH";DF;"D.F"
1530 STOP
1540 END
```

```
;NFF;" FACTORS"
```

FACTOR ANALYSIS OF GENETIC DIVERGENCE IN SESAME

BY

TES P. MATHEW

ABSTRACT OF A THESIS
SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENT FOR THE DEGREE
MASTER OF SCIENCE IN AGRICULTURAL STATISTICS
FACULTY OF AGRICULTURE
KERALA AGRICULTURAL UNIVERSITY

DEPARTMENT OF STATISTICS
COLLEGE OF VETERINARY AND ANIMAL SCIENCES
MANNUTHY - TRICHUR

1988

ABSTRACT

Sesame is an important annual oil seed crop grown in India. It is grown in a very limited area of 1453 hectares in Kerala. The lack of high yielding varieties suitable to the seasons in different regions was the main factor limiting the productivity of sesame in our State. The genetically divergent parents will produce better segregants in the hybridisation programme. The present study was undertaken to delineate the underlying causes of divergence in the sesame plants using the factor analytic methods.

Principal factor and maximum likelihood factor analysis were carried out on a multivariate data on fifteen characters of hundred selected sesame varieties which are grown in upland during rabi, 1981 and rice fallows during summer, 1982. The analysis were done on phenotypic, environment and genotypic correlation matrices under both environments. The phenotypic correlation matrices did not give a stable factor pattern during both rabi and summer seasons. Also the genotypic correlation matrices under the two environments were found not suitable for factor analytic studies. The environment correlation matrices gave stable factor pattern under both the environments and this matrix

was found to be appropriate for factor analytic studies.

The reproductive, height, seed, density and growth factors were identified as the underlying causes of divergence in sesame under the two environments when the Principal Factor Analysis was performed. The Maximum Likelihood Factor Analysis revealed the additional factors viz., quality factor in uplands and a weight factor in rice fellows apart from the above factors. Maximum likelihood method is superior to principal factor analysis method as it gave a better fit of the factor-model. The characters which were most amenable to changes due to selection in these factors were identified in uplands as yield of seeds per plant as reproductive factor, circumference of capsule as seed factor, number of days to maturity as growth factor and number of fruiting nodes per 20 cm as density factor. The same characters were identified as height, growth and density factors in rice fallows. The number of capsules on branches and yield of seeds per plant were identified as reproductive factor, circumference of capsule and number of seeds per capsule as seed factor in rice fallows. The factors relating to growth, productivity and quality were identified as the factors of divergence in sesame.